UNIVERSIDADE FEDERAL DE SÃO CARLOS
CENTRO DE CIÊNCIAS EXATAS E DE TECNOLOGIA
DEPARTAMENTO DE MATEMÁTICA

PEDRO AUGUSTO BUSSOLA

UMA INTRODUÇÃO À CORPOS FINITOS E SUAS APLICAÇÕES

SÃO CARLOS - SP
2025

FUNDAÇÃO UNIVERSIDADE FEDERAL DE SÃO CARLOS
COORDENAÇÃO DOS CURSOS DE GRADUAÇÃO EM MATEMÁTICA (CCM)
Rod. Washington Luís km 235 - SP-310, s/n - Bairro Monjolinho, São Carlos/SP, CEP 13565-905
Telefone: (16) 33518221 - http://www.ufscar.br

DP-TCC-FA nº 11/2026/CCM/CCET

Graduação: Defesa Pública de Trabalho de Conclusão de Curso

Folha Aprovação

FOLHA DE APROVAÇÃO

PEDRO AUGUSTO BUSSOLA

UMA INTRODUÇÃO À CORPOS FINITOS E SUAS APLICAÇÕES

Trabalho de Conclusão de Curso

Universidade Federal de São Carlos – Campus São Carlos

São Carlos, 11 de dezembro de 2025

ASSINATURAS E CIÊNCIAS

Cargo/Função	Nome Completo
Orientador	Humberto Luiz Talpo
Membro da Banca 1	Dimas José Gonçalves
Membro da Banca 2	Pedro Souza Fagundes

Documento assinado eletronicamente por **Humberto Luiz Talpo**, **Professor(a) do Ensino Superior**, em 23/03/2026, às 16:01, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do Decreto nº 8.539, de 8 de outubro de 2015.

Documento assinado eletronicamente por **Pedro Souza Fagundes**, **Professor(a) Adjunto(a)**, em 23/03/2026, às 16:33, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do Decreto nº 8.539, de 8 de outubro de 2015.

Documento assinado eletronicamente por **Dimas Jose Goncalves**, **Professor(a) do Ensino Superior**, em 31/03/2026, às 18:59, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do Decreto nº 8.539, de 8 de outubro de 2015.

A autenticidade deste documento pode ser conferida no site https://sei.ufscar.br/autenticacao, informando o código verificador **2147580** e o código CRC **7098D964**.

PEDRO AUGUSTO BUSSOLA

UMA INTRODUÇÃO À CORPOS FINITOS E SUAS APLICAÇÕES

Monografia apresentada ao curso de Bacharelado em Matemática da Universidade Federal de São Carlos.

Orientador: Prof. Dr. Humberto Luiz Talpo

SÃO CARLOS - SP
2025

Bussola, Pedro Augusto

Uma introdução à corpos finitos e suas aplicações /
Pedro Augusto Bussola -- 2025.
95f.

TCC (Graduação) - Universidade Federal de São Carlos, campus São Carlos, São Carlos
Orientador (a): Humberto Luiz Talpo
Banca Examinadora: Humberto Luiz Talpo, Dimas José Gonçalves, Pedro Souza Fagundes
Bibliografia

1. Corpos finitos. 2. Criptografia. 3. Teoria de códigos. I. Bussola, Pedro Augusto. II. Título.

Ficha catalográfica desenvolvida pela Secretaria Geral de Informática (SIn)

DADOS FORNECIDOS PELO AUTOR

Bibliotecário responsável: Arildo Martins - CRB/8 7180

Dedico este trabalho aos meus pais, que se sacrificaram, lutaram e nunca mediram esforços para me proporcionar as oportunidades que tive. Sem o amor, apoio e dedicação de vocês, esta conquista não seria possível. ...

Agradecimentos

Em primeiro lugar, agradeço profundamente aos meus pais, que sempre foram meu alicerce. O esforço, a dedicação e os inúmeros sacrifícios que fizeram por mim e pela minha educação foram essenciais para que eu pudesse chegar até aqui. Todo este caminho só foi possível graças a vocês.

Agradeço também a todos os professores que fizeram parte da minha trajetória, desde o ensino básico até o ensino superior. Cada um, à sua maneira, contribuiu de forma única para a minha formação acadêmica e pessoal, despertando em mim o interesse pelo conhecimento e o desejo de sempre buscar mais.

Por fim, agradeço aos amigos que construí durante a graduação. Foram vocês que tornaram esses últimos quatro anos mais leves, felizes e inesquecíveis. Cada momento vivido ao lado de vocês ficará para sempre comigo.

A todos, meu muito obrigado!

"O que sabemos é uma gota; o que ignoramos é um oceano."
(Isaac Newton)

Resumo

Esta monografia apresenta um estudo introdutório sobre os corpos finitos, conectando seus fundamentos algébricos às aplicações essenciais na segurança da informação e transmissão de dados. A primeira parte do trabalho estabelece a base teórica, revisando as estruturas de grupos e anéis para construir e caracterizar os corpos finitos como extensões de corpos primos. São detalhados a existência e unicidade de corpos com p^n elementos, a estrutura do grupo multiplicativo e a teoria dos polinômios irredutíveis e ciclotômicos. A segunda parte dedica-se às aplicações práticas destas estruturas. No âmbito da Criptografia, analisam-se os sistemas de chave pública, com ênfase no algoritmo RSA e na complexidade do problema do Logaritmo Discreto. Por fim, apresenta-se uma breve introdução à Teoria de Códigos para a correção de erros, abordando os conceitos de códigos lineares e a caracterização algébrica dos códigos cíclicos, culminando no estudo dos códigos BCH.

Palavras-chave: Corpos finitos, álgebra abstrata, criptografia de chave pública, RSA, logaritmo discreto, teoria de códigos, códigos BCH.

Abstract

This monograph presents an introductory study on finite fields, connecting their algebraic foundations to essential applications in information security and data transmission. The first part of the work establishes the theoretical basis, reviewing the structures of groups and rings to construct and characterize finite fields as extensions of prime fields. The existence and uniqueness of fields with p^n elements, the structure of the multiplicative group, and the theory of irreducible and cyclotomic polynomials are detailed. The second part is dedicated to the practical applications of these structures. In the realm of Cryptography, public-key systems are analyzed, with an emphasis on the RSA algorithm and the complexity of the Discrete Logarithm problem. Finally, a brief introduction to Coding Theory for error correction is presented, addressing the concepts of linear codes and the algebraic characterization of cyclic codes, culminating in the study of BCH codes.

Keywords: Finite fields, abstract algebra, public-key cryptography, RSA, discrete logarithm, coding theory, BCH codes.

Sumário

Introdução

Os corpos finitos são fundamentais na álgebra abstrata e essenciais em áreas como teoria da informação, criptografia e ciência da computação. Este trabalho tem como objetivo apresentar uma introdução abrangente a essas estruturas, conectando seus fundamentos algébricos às aplicações práticas mais relevantes. A motivação para este estudo reside tanto no interesse teórico de caracterizar tais corpos e suas construções explícitas, quanto na sua importância como base para diversas tecnologias contemporâneas.

A fundamentação teórica parte dos conceitos de grupos e anéis para desenvolver a teoria dos corpos finitos, demonstrando a existência e unicidade de corpos com p^n elementos. Enfatiza-se a construção dessas estruturas como extensões algébricas de corpos primos $\mathbb{F}_p$, explorando as propriedades dos polinômios irredutíveis e a natureza cíclica do grupo multiplicativo, um resultado com profundas implicações teóricas e computacionais.

Expandindo o escopo para as aplicações, investiga-se o papel dessas estruturas na segurança e integridade da informação. Em Criptografia, abordam-se desde cifras clássicas até sistemas de chave pública, com destaque para o algoritmo RSA e o problema do Logaritmo Discreto. Na Teoria de Códigos, essencial para a correção de erros, parte-se dos códigos lineares para a caracterização algébrica dos códigos cíclicos e da família de códigos BCH.

A monografia está organizada em quatro capítulos: o primeiro revisa conceitos de grupos, anéis e corpos; o segundo estuda polinômios e extensões; o terceiro dedica-se à teoria dos corpos finitos e suas propriedades; e o quarto aplica esse arcabouço teórico à Criptografia e Teoria de Códigos. O texto pressupõe conhecimentos básicos de álgebra linear e teoria dos números, mantendo-se acessível a estudantes de graduação. As principais referências utilizadas são (LIDL; NIEDERREITER, 1997), (BREšAR, 2019), (LIDL; PILZ, 1997) e (PANARIO, 2007).

1 Conceitos Básicos

Neste capítulo, apresentamos conceitos fundamentais que servirão de base para o desenvolvimento do trabalho. O foco principal será a introdução às estruturas algébricas, como grupos, anéis e corpos, que desempenham um papel central na construção de diversas teorias matemáticas.

A principal referência para esta parte foi: (LIDL; NIEDERREITER, 1997), sobretudo as sessões 1 e 2 do capítulo 1.

1.1 Grupos

Nos conjuntos numéricos, como os inteiros e os reais, é comum utilizarmos as operações de adição e multiplicação. Podemos generalizar essas operações por meio do conceito de operação binária.

Seja G um conjunto. Chamamos de **operação binária** qualquer função do tipo $G \times G \to G$, que associa a cada par ordenado (x, y) um elemento $x * y$ em G.

No entanto, apenas a existência de uma operação binária não é suficiente para garantir propriedades úteis ou estruturas interessantes sobre o conjunto. Por isso, é necessário estabelecer algumas condições adicionais que tornarão esse conjunto mais estruturado e relevante para o nosso estudo. É nesse contexto que surge a definição de **grupo**.

Definição 1. *Um **grupo** é um conjunto G juntamente com uma operação binária $*$ em G que satisfaz as seguintes propriedades:*

1. ***Associatividade**: Para quaisquer $a, b, c \in G$, $a * (b * c) = (a * b) * c$.*
2. ***Elemento identidade**: Existe um elemento $e \in G$, tal que, para todo $a \in G$, $a * e = e * a = a$.*
3. ***Elemento inverso**: Para cada $a \in G$, existe um elemento $a^{-1} \in G$, tal que, $a * a^{-1} = a^{-1} * a = e$.*

*Além disso, um grupo é chamado **abeliano** se vale a comutatividade, ou seja, para todos $a, b \in G$ vale $a * b = b * a$.*

É possível demonstrar que tanto o elemento identidade e quanto o elemento inverso a^{-1} de um dado $a \in G$ são **únicos**, conforme as propriedades anteriormente definidas. Além disso, vale a seguinte relação: para quaisquer $a, b \in G$, temos $(a * b)^{-1} = b^{-1} * a^{-1}$.

Para simplificar a notação, a partir de agora utilizaremos a multiplicação usual para representar a operação no grupo, ou seja, escreveremos ab no lugar de $a * b$. É importante destacar que isso não significa que estamos necessariamente tratando da multiplicação convencional de números. Quando o grupo for *abeliano* (ou comutativo), é comum adotar a notação aditiva, utilizando $a + b$ no lugar de $a * b$ e $-a$ no lugar de a^{-1}.

Graças à propriedade associativa, podemos omitir os parênteses nas composições de múltiplos elementos. Por exemplo, a expressão $a_1 a_2 \cdots a_n$, onde cada $a_j \in G$ e $1 \leq j \leq n$, sempre terá o mesmo valor, independentemente da forma como as operações forem agrupadas.

Além disso, para representar a operação de um elemento $a \in G$ composto consigo mesmo n vezes, com $n \in \mathbb{Z}$, utilizaremos as seguintes convenções:

- **Notação multiplicativa**: $a^n = \underbrace{aa \cdots a}_{n \text{ fatores}}$, chamada de n**-ésima potência** de a.
- **Notação aditiva**: $na = \underbrace{a + a + \cdots + a}_{n \text{ parcelas}}$.

De acordo com a notação adotada, seguem algumas propriedades usuais:

Notação Multiplicativa	**Notação Aditiva**
$a^n a^m = a^{n+m}$	$na + ma = (n+m)a$
$(a^n)^m = a^{nm}$	$m(na) = (mn)a$

Adota-se, por convenção, que para $n = 0 \in \mathbb{N}$, temos $a^0 = e$ na notação multiplicativa e $0a = 0$ na notação aditiva, onde este último "0" representa o elemento identidade do grupo.

Exemplo 2. ***(i)*** *Note que números inteiros com a operação de adição é um grupo, com o elemento identidade* 0*, e o inverso de um inteiro* a *é o inteiro* $-a$*. Denotamos este grupo por* $(\mathbb{Z}, +)$ *.*

(ii) *O conjunto consistindo de um único elemento* e*, com a operação* $*$ *definida por* $e * e = e$*, forma um grupo. Chamamos de grupo trivial.*

Definição 3. *Um grupo* G *é dito* ***finito*** *se possui um número finito de elementos. Esse número é chamado de ordem do grupo, e denotamos por* $|G|$*.*

Quando um grupo não é finito, chamamos de grupo ***infinito****.*

Neste momento, preparamos o cenário para um exemplo muito importante de grupo finito. Sejam S um conjunto e $R \subseteq S \times S$, R é chamada de **relação de equivalência** em S se satisfaz:

1. **Reflexividade**: $(s, s) \in R$ para todo $s \in S$.
2. **Simetria**: Se $(s, t) \in R$, então $(t, s) \in R$.
3. **Transitividade**: Se $(s, t) \in R$ e $(t, u) \in R$, então $(s, u) \in R$.

Com essa relação de equivalência R em S e um elemento $s \in S$, a **classe de equivalência** de s é o conjunto:

$$[s] := \{t \in S \mid (s, t) \in R\}.$$

Exemplo 4. *Em $\mathbb{Z}$, a relação "congruência módulo n" definida por $a \equiv b \mod n \Leftrightarrow a = b + kn$ para algum $k \in \mathbb{Z}$ é:*

- *Reflexiva: $a \equiv a \mod n$*
- *Simétrica: $a \equiv b \mod n \Rightarrow b \equiv a \mod n$*
- *Transitiva: $a \equiv b \mod n$ e $b \equiv c \mod n \Rightarrow a \equiv c \mod n$*

Suas classes são $[k] = \{k + mn \mid m \in \mathbb{Z}\}$ ***para*** $0 \leq k < n$**.**

Com isso, o conjunto formado por essas classes de equivalência módulo n é $\{[0], [1], \cdots, [n-1]\}$, e com a operação definida por $[a] + [b] = [a + b]$ é um grupo que denotamos por $\mathbb{Z}_n$. Esse grupo é de extrema importância para resultados futuros.

Definição 5. *Seja G um grupo, dizemos que $S \subseteq G$ é um* ***subgrupo*** *se S, com a operação de G, também for um grupo. Um subgrupo de G que não seja o subgrupo trivial $\{e\}$ e G, é chamado de* ***subgrupo não trivial****.*

Definição 6. *Para qualquer $a \in G$, o* ***subgrupo gerado por*** *a é:*

$$\langle a \rangle := \{a^n \mid n \in \mathbb{Z}\}.$$

Se $\langle a \rangle$ é finito, então a ordem do grupo é a ordem de a. Caso seja infinito, dizemos que a tem ordem infinita. Podemos também chamar $\langle a \rangle$ de grupo cíclico gerado por a, ou somente grupo cíclico.

Seja G um grupo e H um subgrupo de G. Podemos definir uma relação de equivalência R_H em G por:

$$(a, b) \in R_H \iff a^{-1}b \in H.$$

Essa relação é conhecida como **congruência à esquerda módulo** H, e as classes de equivalência induzidas por ela são denominadas **classes laterais à esquerda** de H em G.

Definição 7. *Dada $a \in G$, a* ***classe lateral à esquerda*** *de H por a é o subconjunto*

$$aH := \{ah \mid h \in H\}.$$

E chamamos ***classe lateral à direita*** *de H por a o subconjunto*

$$Ha := \{ha \mid h \in H\}.$$

Observação 8. *Essas classes formam uma partição de G, ou seja, cada elemento de G pertence a exatamente uma classe lateral.*

Como consequência das definições anteriores, conclui-se que o conjunto G é a união disjunta de suas classes laterais à esquerda em relação a H.

Definição 9. *O número de classes laterais à esquerda de H em G é chamado de* ***índice de H em G****, denotado por $[G : H]$.*

Teorema 10 (Teorema de Lagrange)**.** *Se G é um grupo finito e H é um subgrupo de G, então*

$$|G| = [G : H] \cdot |H|.$$

Em particular, a ordem de qualquer subgrupo H de G divide a ordem de G. Ademais, a ordem de qualquer elemento $a \in G$ divide $|G|$.

Demonstração. Omitiremos essa demonstração, porém pode ser encontrada em (BREšAR, 2019) página 137. □

Teorema 11. *Seja $\langle a \rangle$ um grupo cíclico de ordem m. Então:*

1. *Todo subgrupo de um grupo cíclico é também cíclico.*
2. *Para todo $k \in \mathbb{Z}$, o elemento a^k gera um subgrupo de ordem $\dfrac{m}{mdc(k,m)}$.*
3. *Para cada divisor positivo d de m, existe um único subgrupo de $\langle a \rangle$ de ordem d, e, consequentemente, um único subgrupo de índice m/d.*
4. *Para cada divisor positivo d de m, o grupo $\langle a \rangle$ possui exatamente $\varphi(d)$ elementos de ordem d, onde φ é a função totiente de Euler.*
5. *O grupo $\langle a \rangle$ possui exatamente $\varphi(m)$ geradores, ou seja, elementos a^r tais que $\langle a^r \rangle = \langle a \rangle$. Esses geradores são os elementos a^r com $mdc(r, m) = 1$.*

Demonstração. 1. Seja H um subgrupo de $\langle a \rangle$, com $H \neq \{e\}$. Como $\langle a \rangle$ é cíclico, existe algum $a^n \in H$ com $n > 0$. Seja d o menor inteiro positivo tal que $a^d \in H$. Mostraremos que $H = \langle a^d \rangle$. Seja $a^s \in H$. Pela divisão euclidiana, existem inteiros q

e r com $0 \leq r < d$ e $s = qd + r$. Assim, $a^r = a^s(a^{-d})^q \in H$, pois $a^s \in H$ e $a^{-d} \in H$ implicam $a^r \in H$. Pela minimalidade de d, concluímos que $r = 0$, ou seja, $d \mid s$, e portanto $a^s \in \langle a^d \rangle$. Logo, $H \subseteq \langle a^d \rangle$. Como $a^d \in H$, também $\langle a^d \rangle \subseteq H$, e assim $H = \langle a^d \rangle$, ou seja, H é cíclico.

2. Seja $d = mdc(k, m)$. A ordem do elemento a^k é o menor inteiro $n > 0$ tal que $(a^k)^n = a^{kn} = e$. Isso ocorre se, e somente se, $m \mid kn$. Como $mdc(k, m) = d$, temos $k = dk'$, $m = dm'$ com $mdc(k', m') = 1$, e assim $m \mid kn \Leftrightarrow dm' \mid dk'n \Leftrightarrow m' \mid k'n$. Como $mdc(k', m') = 1$, o menor n tal que $m' \mid k'n$ é $n = m'$, ou seja, a ordem de a^k é m/d.

3. Seja d um divisor positivo de m. Pelo item (2), o elemento $a^{m/d}$ tem ordem d, e o subgrupo que ele gera tem ordem d. Para mostrar a unicidade, suponha que a^k também gere um subgrupo de ordem d. Então, pelo item (2), $mdc(k, m) = m/d$, o que implica que a^k pertence a $\langle a^{m/d} \rangle$. Como ambos geram subgrupos de mesma ordem, temos $\langle a^k \rangle = \langle a^{m/d} \rangle$.

4. Um elemento a^k tem ordem d se, e somente se, $mdc(k, m) = m/d$. Escrevendo $k = mh/d$, temos que $mdc(h, d) = 1$. Assim, o número de tais k é igual ao número de inteiros h com $1 \leq h < d$ e $mdc(h, d) = 1$, ou seja, $\varphi(d)$.

5. Um elemento a^r é gerador de $\langle a \rangle$ se, e somente se, a ordem de a^r é m. Pelo item (2), isso ocorre se, e somente se, $mdc(r, m) = 1$. Assim, existem $\varphi(m)$ geradores no total.

□

Definição 12. *Sejam G e G' grupos. Uma função $\varphi : G \to G'$ é um* ***homomorfismo*** *se, para todo $x, y \in G$, temos:*

$$\varphi(xy) = \varphi(x)\varphi(y).$$

No caso particular quando um homomorfismo é bijetivo chamamos de ***isomorfismo****. Denotamos que $G \cong G'$ se existe um isomorfismo entre eles. E quando $G = G'$ chamamos φ de* ***automorfismo****.*

Definição 13. *O* ***núcleo*** *de um homomorfismo φ é o conjunto:*

$$\ker(\varphi) := \{x \in G \mid \varphi(x) = e\},$$

e sua ***imagem*** *é:*

$$\mathrm{im}(\varphi) := \{\varphi(x) \mid x \in G\}.$$

Vale destacar algumas propriedades importantes dos homomorfismos de grupos.

Seja $\varphi : G \to G'$ um homomorfismo. Assim como seria de se esperar, φ preserva as principais propriedades do grupo: a identidade e os inversos. Ou seja,

$$\varphi(e) = e' \quad \text{e} \quad \varphi(a^{-1}) = \varphi(a)^{-1}, \quad \text{para todo } a \in G.$$

Além disso, o kernel é fechado por conjugação, o que quer dizer que, se $k \in \ker(\varphi)$ e $x \in G$, então também temos $xkx^{-1} \in \ker(\varphi)$.

Por fim, uma propriedade essencial: φ é injetiva se, e somente se, seu núcleo é trivial, isto é,

$$\varphi \text{ é injetiva } \iff \ker(\varphi) = \{e\}.$$

Definição 14. *Seja H um subgrupo de um grupo G. Dizemos que H é um **subgrupo normal** de G se, para todo $a \in G$ e todo $h \in H$, temos:*

$$aha^{-1} \in H.$$

Ou seja, H é estável sob conjugação por elementos de G.

Chamamos de **automorfismo interno** qualquer automorfismo que surge da conjugação.

Teorema 15. *Seja H um subgrupo de G. As seguintes afirmações são equivalentes:*

1. *H é normal em G;*
2. *H é igual a todos os seus conjugados, ou seja, $aHa^{-1} = H$ para todo $a \in G$;*
3. *H é invariante sob todos os automorfismos internos de G;*
4. *Toda classe lateral à esquerda de H coincide com a correspondente classe lateral à direita, isto é,*
$$aH = Ha \quad \text{para todo } a \in G.$$

Teorema 16. *Seja H um subgrupo normal de um grupo G. Então, o conjunto das classes laterais à esquerda de H em G forma um grupo com a operação:*

$$(aH)(bH) := (ab)H.$$

Definição 17. *Seja H um subgrupo normal de um grupo G. O grupo formado pelas classes laterais à esquerda de H em G, com a operação definida por*

$$(aH)(bH) := (ab)H,$$

*é chamado de **grupo quociente** (ou **fator**) de G por H e é denotado por G/H.*

Se G é um grupo finito e $H \trianglelefteq G$ (ou seja, H é normal em G), então o grupo quociente G/H também é finito. Além disso, sua ordem é igual ao índice de H em G:

$$|G/H| = [G : H] = \frac{|G|}{|H|}.$$

Essa relação é consequência direta do Teorema 10.

Teorema 18 (Teorema do Homomorfismo)**.** *Seja $\varphi : G \to G_1$ um homomorfismo de grupos. Então, o núcleo $\ker(\varphi)$ é um subgrupo normal de G, e se φ é sobrejetor temos que G_1 é isomorfo ao grupo quociente $G/\ker(\varphi)$.*

Reciprocamente, se H é um subgrupo normal de G, então a aplicação

$$\psi : G \to G/H, \quad \psi(a) := aH,$$

é um homomorfismo sobrejetivo de G em G/H, cujo núcleo é exatamente H.

Demonstração. Omitiremos essa demonstração, porém pode ser encontrada em (BREšAR, 2019) página 153, denotado por outro nome "Isomorphism Theorem".

□

Definição 19. *Seja S um subconjunto não vazio de um grupo G. O **normalizador** de S em G é definido por:*

$$N(S) := \{a \in G \mid aSa^{-1} = S\}.$$

Teorema 20. *Para qualquer subconjunto não vazio S de um grupo G, o normalizador $N(S)$ é um subgrupo de G. Além disso, há uma correspondência biunívoca entre as classes laterais à esquerda de G módulo $N(S)$ e os distintos conjugados de S da forma aSa^{-1}.*

Demonstração. Temos $e \in N(S)$, e se $a, b \in N(S)$, então $a^{-1} \in N(S)$ e $ab \in N(S)$. Logo, $N(S)$ é subgrupo de G.

Seja $C_L = \{aN(S) \mid a \in G\}$ o conjunto das classes laterais à esquerda de G módulo $N(S)$ e seja $\overline{S} = \{aSa^{-1} \mid a \in G\}$ o conjunto dos distintos conjugados de S .

Consideramos a aplicação $\phi : C_L \to \overline{S}$ definida por:

$$\phi(aN(S)) = aSa^{-1}$$

Devemos garantir que a imagem de ϕ não depende do representante a e que conjugados distintos correspondem a classes laterais distintas. Se $aSa^{-1} = bSb^{-1}$, então multiplicando à esquerda por a^{-1} e à direita por a, obtemos:

$$S = a^{-1}(bSb^{-1})a = (a^{-1}b)S(b^{-1}a) = (a^{-1}b)S(a^{-1}b)^{-1}$$

Isto implica que $a^{-1}b \in N(S)$, o que é a condição necessária e suficiente para que a e b pertençam à mesma classe lateral à esquerda de $N(S)$, ou seja, $aN(S) = bN(S)$. Portanto, ϕ é bem-definida e injetora.

Dado qualquer elemento $Y \in \overline{S}$, Y é, por definição, um conjugado de S, logo Y é da forma aSa^{-1} para algum $a \in G$. O elemento $aN(S) \in C_L$ é tal que $\phi(aN(S)) = aSa^{-1} = Y$. Como todo elemento em $\overline{S}$ é a imagem de alguma classe lateral em C_L, a aplicação ϕ é sobretiva.

Uma vez que ϕ é injetora e sobretiva, ela é uma correspondência biunívoca entre as classes laterais à esquerda de G módulo $N(S)$ e os distintos conjugados de S. □

Definição 21. *O* ***centro*** *de um grupo G é o conjunto:*

$$C := \{c \in G \mid ac = ca \text{ para todo } a \in G\}.$$

Teorema 22 (Equação das classes)**.** *Seja G um grupo finito com centro C. Então:*

$$|G| = |C| + n_1 + n_2 + \cdots + n_k,$$

onde cada $n_i \geq 2$ e divide $|G|$. Os termos $n_1, n_2, \ldots, n_k$ correspondem aos tamanhos das classes de conjugação não triviais de G.

Demonstração. A relação de conjugação em G é uma relação de equivalência, e suas classes de equivalência são as classes de conjugação. Assim, a soma dos tamanhos de todas as classes de conjugação é igual a $|G|$.

Os elementos de C são exatamente aqueles que formam classes de conjugação com apenas um elemento (pois comutam com todos). Os demais elementos pertencem a classes de conjugação com dois ou mais elementos.

Cada classe de conjugação $a^G = \{gag^{-1} \mid g \in G\}$ tem tamanho igual ao número de classes laterais de G módulo $N(\{a\})$, ou seja, $[G : N(\{a\})]$, que divide $|G|$.

Logo, temos a fórmula:

$$|G| = |C| + \sum_{i=1}^{k} [G : N(\{a_i\})] = |C| + n_1 + \cdots + n_k.$$

□

1.2 Anéis e Corpos

Definição 23. *Um* anel $(R, +, \cdot)$ *é um conjunto R munido com duas operações binárias, denotadas por $+$ e $\cdot$, tal que*

1. *o conjunto R é um grupo abeliano com respeito a $+$;*
2. *a operação $\cdot$ é associativa, isto é, $(a \cdot b) \cdot c = a \cdot (b \cdot c)$, para quaisquer $a, b, c \in R$; e*
3. *para todos $a, b, c \in R$ vale a distributiva, isto é,*

$$a \cdot (b + c) = a \cdot b + a \cdot c \text{ e } (b + c) \cdot a = b \cdot a + c \cdot a\,.$$

Vamos aqui utilizar R como o anel $(R, +, \cdot)$ e reforçar que as operações $+$ e $\cdot$ não são necessariamente as operações conhecidas utilizadas com números.

Usaremos 0 (chamado de *elemento zero*) para denotar o elemento neutro do grupo abeliano R com respeito à adição e $-a$ para denotar o oposto de a. Além disso, $a + (-b)$ será escrito como $a - b$ e $a \cdot b$ será denotado simplesmente por ab.

Temos como consequência da definição de anel que, para todo $a \in R$, $a0 = 0a = 0$. Isso implica que $(-a)b = a(-b) = -ab$, para quaisquer $a, b \in R$.

Definição 24. *Um anel R associativo é dito um:*

1. anel com identidade *se R tem identidade multiplicativa, ou seja, existe um elemento $e \in R$ tal que, para todo $a \in R$, $ae = ea = a$;*
2. comutativo, *se $\cdot$ é comutativa;*
3. domínio de integridade *se é um anel comutativo com identidade $e \neq 0$ onde $ab = 0$ implica que $a = 0$ ou $b = 0$;*
4. anel de divisão *(ou* corpo não comutativo*) se, sobre $\cdot$, os elementos não nulos de R formam um grupo;*
5. *Um anel de divisão comutativo é dito um* corpo.

Exemplo 25. *O conjunto dos inteiros $\mathbb{Z}$ com as operações usuais é um domínio de integridade, mas não um corpo.*

Teorema 26. *Todo domínio de integridade finito é um corpo.*

Demonstração. Sejam $a_1, a_2, \ldots, a_n$ os elementos de um domínio de integridade finito R. Fixemos um elemento não nulo $a \in R$ e consideremos os produtos $aa_1, aa_2, \ldots, aa_n$. Tais produtos são distintos, pois se $aa_i = aa_j$, então $a(a_i - a_j) = 0$. Como $a \neq 0$ é o caso que $a_i - a_j = 0$ ou $a_i = a_j$. Logo, os elementos de R são da forma aa_i. Em particular, existe i, com $1 \leq i \leq n$, tal que $e = aa_i$, onde e é a identidade em R. Como R é comutativo, temos $a_i a = e$ e, daí, a_i é o inverso multiplicativo de a. Portanto, os elementos não nulos de R formam um grupo abeliano e R é um corpo. □

Definição 27. *Um subconjunto S de um anel R é dito um* subanel *de R se S é fechado nas operações e forma um anel sobre as operações de R.*

Definição 28. *Um subconjunto J de um anel R é dito um* ideal *quando J é um subanel de R e para todo $a \in J$ e $r \in R$ temos ra e ar pertencendo a J.*

Definição 29. *Seja R um anel comutativo. Um ideal J de R é* principal *se existir um elemento $a \in R$ tal que J seja gerado,como ideal, por a, denotando por $J = \langle a \rangle$. Neste caso, J é também chamado de* ideal principal gerado por a.

Como ideais são subgrupos normais do grupo aditivo de um anel, segue de imediato que um ideal J de um anel R define uma partição de R em classes laterais disjuntas, chamadas *classes de congruência* módulo J.

Elementos $a, b \in R$ são ditos *congruentes* módulo J se eles estão na mesma classe de congruência módulo J ou, de forma equivalente, $a - b \in J$.

Definição 30. *O anel das classes de congruência do anel R módulo o ideal J sobre as operações*

$$(a + J) + (b + J) = (a + b) + J \quad e \quad (a + J)(b + J) = ab + J,$$

é chamado anel quociente *de R módulo J e é denotado por R/J.*

Exemplo 31 (O anel quociente $\mathbb{Z}/\langle n\rangle$)**.** *Como no caso de grupos, vamos denotar a classe lateral ou classe de congruência do inteiro a módulo o inteiro positivo n por $[a]$ ou $a + \langle n\rangle$, onde $\langle n\rangle$ é o ideal principal gerado por n. Os elementos em $\mathbb{Z}/\langle n\rangle$ são*

$$[0] = 0 + \langle n\rangle, [1] = 1 + \langle n\rangle, \ldots, [n - 1] = n - 1 + \langle n\rangle.$$

Para o caso que $n = p$ com p primo, temos que $\mathbb{Z}/\langle p\rangle$ é um corpo.

De fato, pelo Teorema 26, basta mostrar que $\mathbb{Z}/\langle p\rangle$ é um domínio de integridade. Note que $[1]$ é o elemento identidade de $\mathbb{Z}/\langle p\rangle$, e que $[a][b] = [ab] = [0]$ se, e somente se, $ab = kp$ para algum inteiro k. Mas, como p é primo, p divide ab se, e somente se, p divide pelo menos um dos fatores. Portanto, ou $[a] = [0]$ ou $[b] = [0]$, o que mostra que $\mathbb{Z}/(p)$ não possui divisores de zero. □

Existe uma extensão natural de grupos para anéis na definição de homomorfismo. Uma função $\varphi\colon R \to S$ de um anel R a um anel S é dita um *homomorfismo de anéis* se, para todo $a, b \in R$ temos

$$\varphi(a + b) = \varphi(a) + \varphi(b) \quad \text{e} \quad \varphi(ab) = \varphi(a)\varphi(b).$$

Logo, um homomorfismo $\varphi\colon R \to S$ preserva as duas operações em R e induz um homomorfismo do grupo aditivo R ao grupo aditivo S. Além disso, o conjunto

$$\ker \varphi = \{a \in R \mid \varphi(a) = 0 \in S\}$$

é chamado *núcleo* (ou *kernel*) de φ. O conceito de *isomorfismo de anéis* é analógo ao que temos na Definição 12. O Teorema do homomorfismo para anéis é similar ao que vimos para grupos.

Teorema 32 (Teorema do Homomorfismo para Anéis)**.** *Se φ é um homomorfismo de anéis sobrejetivo de um anel R a um anel S, então* $\ker \varphi$ *é um ideal de R. Além disso, o anel S é isomorfo ao anel quociente $R/\ker \varphi$. Por outro lado, Se J é um ideal do anel R, então a função $\psi\colon R \to R/J$ dada por $\psi(a) = a + J$ é um homomorfismo sobrejetivo de R a R/J com núcleo J.*

Definição 33. *Dado um número primo p, seja $\mathbb{F}_p$ o conjunto de inteiros $\{0, 1, \ldots, p-1\}$ e tomemos $\varphi\colon \mathbb{Z}/\langle p\rangle \to \mathbb{F}_p$ a função definida por $\varphi([a]) = a$, com $a = 0, 1, \ldots, p-1$. Então $\mathbb{F}_p$, com a estrutura de corpo induzida por φ é um corpo finito, chamado* corpo finito de ordem p *(ou corpo de Galois de ordem p).*

Definição 34. *Seja R um anel qualquer. Se existir um inteiro positivo p tal que $pr = 0$ para todo $r \in R$, então o menor inteiro com essa propriedade é dito a* característica R *e dizemos que R tem característica p positiva. Se tal inteiro positivo não existir, dizemos que R é de característica* 0*.*

Teorema 35. *Seja $R \neq \{0\}$ um anel de característica p positiva, com identidade e sem divisores de zero. Nestas condições, a característica p é prima.*

Demonstração. Como R contém elementos não nulos sua característica é $n \geq 2$. Suponhamos que n não seja primo, então podemos escrever $n = km$, com $k, m \in \mathbb{Z}$ e $1 < k, m < n$. Daí $0 = ne = (km)e = (ke)(me)$ implicando que $ke = 0$ ou $me = 0$, pois R não tem divisores de zero. Segue que $kr = (ke)r = 0$ ou $mr = (me)r = 0$ para todo $r \in R$, contradizendo a definição da característica de n. □

Corolário 36. *Todo corpo finito é de característica prima.*

Demonstração. Pelo Teorema 35 é suficiente mostrarmos que um corpo finito F tem característica positiva. Vamos considerar os múltiplos $e, 2e, 3e, \ldots$ da identidade. Como F tem um número finito de elementos distintos, existem inteiros k e m com $1 \leq k < m$ tais que $ke = me$, isto é, $(m-k)e = 0$. Segue que F tem característica positiva. □

Teorema 37. *Seja R um anel comutativo de característica prima p. Então*

$$(a+b)^{p^n} = a^{p^n} + b^{p^n} \qquad e \qquad (a-b)^{p^n} = a^{p^n} - b^{p^n}\,,$$

para $a, b \in R$ e $n \in \mathbb{N}$ quaisquer.

Demonstração. Vamos utilizar do seguinte fato

$$\binom{p}{i} = \frac{p(p-1)\cdots(p-i+1)}{1\cdot 2\cdots i} \equiv 0 \bmod p\,,$$

para todo $i \in \mathbb{Z}$, com $0 < i < p$. Notemos que $\binom{p}{i}$ é um inteiro e que o p do numerador não pode ser cancelado. Daí, pelo binômio de Newton

$$(a+b)^p = a^p + \binom{p}{1}a^{p-1}b + \cdots + \binom{p}{p-1}ab^{p-1} + b^p = a^p + b^p\,,$$

e por indução em n segue a primeira identidade. Pelo que já mostramos, temos

$$a^{p^n} = \big((a-b)+b\big)^{p^n} = (a-b)^{p^n} + b^{p^n}\,,$$

e segue a segunda identidade. □

Seja R um anel comutativo com identidade. Um elemento $a \in R$ é chamado de *divisor* de $b \in R$ se existe $c \in R$ tal que $ac = b$. Uma *unidade* de R é um divisor da identidade; dois elementos $a, b \in R$ são ditos *associados* se existe uma unidade ϵ de R tal que $a = b\epsilon$.

Um elemento $p \in R$ com R um anel comutativo, é chamado de *elemento primo* se $p \neq 0$, p não é uma unidade e se, sempre que p divide um produto ab (com $a, b \in R$), então p divide a ou p divide b. Um ideal $P \neq R$ do anel R é chamado de *ideal primo* se para quaisquer $a, b \in R$ temos que $ab \in P$ implica $a \in P$ ou $b \in P$. Um ideal $M \neq R$ de R é chamado de *ideal maximal* se para qualquer ideal J de R a propriedade $M \subseteq J$ implica $J = R$ ou $J = M$.

Além disso, R é dito um *domínio de ideais principais* se R é um domínio de integridade e se todo ideal J de R é principal - isto é, existe um elemento gerador a para J tal que $J = \langle a \rangle = \{ra \mid r \in R\}$.

Teorema 38. *Seja R um anel comutativo com identidade. Então:*

1. *um ideal M de R é um ideal maximal se e somente se o quociente R/M é um corpo;*
2. *um ideal P de R é um ideal primo se e somente se o quociente R/P é um domínio de integridade;*
3. *todo ideal maximal de R é um ideal primo;*
4. *se R é um domínio de ideais principais, então $R/\langle c \rangle$ é um corpo se e somente se $c \neq 0$ é um elemento primo de R.*

Demonstração. Item 1. Seja M um ideal maximal de R. Para $a \in R$, com $a \notin M$, o conjunto $J = \{ar + m \mid r \in R, m \in M\}$ é um ideal próprio de R contendo M. Pela maximalidade de M, segue que $J = R$. Em particular, podemos tomar elementos $r \in R$ e $m \in M$ de forma que $ar + m = 1$, onde 1 é a identidade multiplicativa de R. Ou seja, se $a + M \neq 0 + M$ é um elemento de R/M diferente do seu elemento neutro, então ele possui inverso multiplicativo. Com efeito,

$$(a + M)(r + M) = ar + M = (1 - m) + M = 1 + M.$$

Logo, R/M é um corpo. Por outro lado, sejam R/M um corpo e J um ideal de R tal que M está propriamente contido em J. Então, dado a em J, mas não em M, a classe de congruência $a+M$ possui inverso multiplicativo de modo que $(a+M)(r+M) = 1+M$, para algum $r \in R$. Então, existe $m \in M$ tal que $ar + m = 1$. Como J é um ideal, temos $1 \in J$ e daí $\langle 1 \rangle = R \subseteq J$. Portanto, $J = R$ e M é um ideal maximal de R.

Item 2. Seja P um ideal primo de R. Então, R/P é um anel comutativo com identidade $1 + P \neq 0 + P$. Tomemos $(a + P)(b + P) = 0 + P$, daí $ab \in P$. Como P é um ideal primo,

ou $a \in P$ ou $b \in P$, isto é, ou $a + P = 0 + p$ ou $b + P = 0 + P$. Segue que R/P não tem divisores de zero e é, portanto, um domínio de integridade. A volta é análoga.

Item 3. Segue dos itens 1 e 2, pois todo corpo é um domínio de integridade.

Item 4. Seja $c \in R$. Se c é uma unidade, então $\langle c \rangle = R$ e o anel $R/\langle c \rangle$ tem somente um elemento e, logo, não é um corpo. Se c não é uma unidade ou um elemento primo, então c tem um divisor $a \in R$ que não é uma unidade ou um elemento associado de c. Notemos que a é não nulo, senão teríamos $c = 0$ e a seria um elemento associado de c. Podemos, então, escrever $c = ab$, com $b \in R$. Afirmamos que $a \notin \langle c \rangle$, pois se estivesse, existiria $d \in R$ tal que $a = cd = abd$ ou $a(1 - bd) = 0$. Como $a \neq 0$, teríamos que $bd = 1$ e d seria uma unidade, contradizendo o fato que a não é um associado de c. Segue que $\langle c \rangle \subset \langle a \rangle \subset R$ sao contenções próprias. Daí, $R/\langle c \rangle$ não pode ser um corpo pelo item 1. Por fim, temos o caso em que c é um elemento primo. Temos que $\langle c \rangle \neq R$, pois c não pode ser uma unidade. Além disso, se J é um ideal de R contendo $\langle c \rangle$, então, como R é um domínio de ideais principais, existe $a \in R$ tal que $J = \langle a \rangle$. Consequentemente, $c \in \langle a \rangle$ e a é um divisor de c. Com isso, vemos que a é uma unidade ou um associado de c. Daí, ou $J = R$ ou $J = \langle c \rangle$ e concluímos que $\langle c \rangle$ é um ideal maximal de R. Portanto, $R/\langle c \rangle$ é um corpo pelo item 1. □

2 Polinômios e Extensões de Corpos

2.1 Polinômios

Seja R um anel arbitrário. Definimos um polinômio na variável x sobre R (com coeficientes em R) como uma soma infinita

$$\sum_{i=0}^{\infty} a_i x^i = a_0 + a_1 x + a_2 x^2 + \cdots$$

onde $a_i \in R$ e todos, exceto uma quantidade finita de $a_i's$, são 0. Os elementos a_i são chamados de coeficientes, e x é um símbolo formal chamado de variável (ou indeterminada). O coeficiente a_0 é chamado de termo constante. Se n é tal que $a_i = 0$ para todo $i > n$ e $a_n \neq 0$, então usualmente escrevemos o polinômio acima como

$$a_0 + a_1 x + \cdots + a_n x^n.$$

Chamamos $R[x]$ o conjunto de todos os polinômios sobre R. O elemento nulo de $R[x]$ é o polinômio que tem todos os coeficientes nulos. Esse polinômio é chamado de *polinômio nulo* e é denotado por 0. Deve sempre estar claro pelo contexto se 0 representa a identidade em R ou o polinômio nulo.

Definição 39. *Seja $f(x) = \sum_{i=0}^{n} a_i x^i$ um polinômio sobre R diferente do polinômio nulo, de modo que possamos supor $a_n \neq 0$. Dizemos que a_n é o* coeficiente líder *de $f(x)$ e n é o* grau *de $f(x)$, denotado por $n = \deg\big(f(x)\big) = \deg(f)$.*

Por convenção, definimos $\deg(0) = -\infty$. Polinômios de grau 0 são chamados de *polinômios constantes*. Se R possui identidade 1 e se o coeficiente líder de $f(x)$ é 1, então $f(x)$ é dito um *polinômio mônico*.

Para definir as operações em $R[x]$, considere dois polinômios $f(x) = \sum_{i=0}^{\infty} a_i x^i$ e $g(x) = \sum_{i=0}^{\infty} b_i x^i$, onde apenas um número finito de coeficientes é não nulo. A adição e a multiplicação são definidas por:

$$f(x) + g(x) = \sum_{i=0}^{\infty} (a_i + b_i) x^i$$

$$f(x) \cdot g(x) = \sum_{k=0}^{\infty} c_k x^k, \quad \text{onde } c_k = \sum_{i+j=k} a_i b_j$$

O coeficiente c_k é obtido pela convolução dos coeficientes de f e g. Note que uma vez que R é um anel, verifica-se que estas operações satisfazem os axiomas de anel

(associatividade, distributividade, existência de neutro aditivo, etc.). Portanto, o conjunto $R[x]$ munido destas operações usuais constitui um anel, denominado anel de polinômios sobre R.

Teorema 40. *Sejam $f, g \in R[x]$. Então*

$$\deg(f+g) \leq \max\big(\deg(f), \deg(g)\big) \quad e \quad \deg(fg) \leq \deg(f) + \deg(g).$$

Se R for um domínio de integridade, temos

$$\deg(fg) = \deg(f) + \deg(g).$$

Teorema 41. *Seja R um anel. Então:*

1. *$R[x]$ é comutativo se e somente se R é comutativo;*
2. *$R[x]$ é um anel com unidade se e somente se R tem unidade;*
3. *$R[x]$ é um domínio de integridade se e somente se R é um domínio de integridade.*

Seja F um corpo. O conceito de divisibilidade, quando aplicado ao anel $F[x]$, nos leva à seguinte definição.

Definição 42. *Sejam $f, g \in F[x]$. Dizemos que o polinômio g* divide *f se existir um polinômio $h \in F[x]$ tal que $f = gh$. Podemos dizer que g é um* divisor *de f; f é um* múltiplo *de g; ou f é* divisível *por g.*

Teorema 43 (Algoritmo da Divisão)**.** *Seja $g \neq 0$ um polinômio em $F[X]$. Então, para todo $f \in F[x]$ existem polinômios $q, r \in F[x]$ tais que*

$$f = qg + r, \quad onde \quad \deg r < \deg g\,.$$

Teorema 44. *O anel $F[x]$ é um domínio de ideais principais. Mais precisamente, para todo ideal $J \neq \langle 0\rangle$ de $F[x]$ existe um único polinômio mônico $g \in F[x]$ com $J = \langle g\rangle$.*

Demonstração. Pelo item 3 do Teorema 41 o anel $F[x]$ é um domínio de integridade. Suponhamos que $J \neq \langle 0\rangle$ é um ideal de $F[x]$. Sejam $h(x)$ um polinômio não nulo de menor grau contido em J e b o coeficiente líder de $h(x)$. Além disso, tomemos $g(x) = b^{-1}h(x)$. Então, $g \in J$ e g é mônico. Tomando $f \in J$ qualquer, o algoritmo da divisão nos dá $q, r \in F[x]$ com $f = qg + r$ e $\deg(r) < \deg(g) = \deg(h)$. Como J é um ideal, temos $f - qg = r \in J$ e, pela definição de h, o resto $r = 0$. Logo, f é um múltiplo de g e $J = \langle g\rangle$. Se $g_1 \in F[x]$ é outro polinômio mônico tal que $J = \langle g_1\rangle$, então $g = c_1 g_1$ e $g_1 = c_2 g$, com $c_1, c_2 \in F[x]$. Ora, isso implica em $g = c_1 c_2 g$ e daí $c_1 c_2 = 1$. Segue que c_1 e c_2 são polinômios constantes. Como g e g_1 são ambos polinômios mônicos, temos $g = g_1$, como queríamos. □

Teorema 45. *Sejam $f_1, \ldots, f_n$ polinômios não todos nulos em $F[x]$. Então, existe um polinômio mônico unicamente determinado $d \in F[x]$ com as seguintes propriedades:*

1. *d divide cada f_j, com $1 \leq j \leq n$;*
2. *se $c \in F[x]$ divide cada f_j, com $1 \leq j \leq n$, então c divide d.*

Além disso, d pode ser escrito na forma

$$d = b_1 f_1 + \cdots + b_n f_n \qquad \text{onde } b_1, \ldots b_n \in F[x]\,. \tag{2.1}$$

Demonstração. Tomemos o conjunto J dos polinômios da forma

$$c_1 f_1 + \cdots + c_n f_n, \quad \text{com} \quad c_1, \ldots, c_n \in F[x].$$

Temos que J é um ideal de $F[x]$. Como existe $f_j \neq 0$, temos $J \neq \langle 0 \rangle$ e o Teorema 44 nos diz que existe um polinômio mônico $d \in F[x]$ tal que $J = \langle d \rangle$. O item 1 e a expressão (2.1) seguem da construção de d; a propriedade em 2 segue de (2.1). Por fim, suponhamos que d_1 seja outro polinômio mônico em $F[x]$ satisfazendo as duas propriedades do enunciado. Então, d e d_1 são divisiveis um pelo outro, daí $\langle d \rangle = \langle d_1 \rangle$. Aplicando a parte da unicidade apresentada no Teorema 44 nos dá $d = d_1$. □

O polinômio mônico d que aparece no Teorema anterior é dito o *máximo divisor comum* dos $f_1, \ldots, f_n$ e é denotado por $d = mdc(f_1, \ldots, f_n)$. Se $mdc(f_1, \ldots, f_n) = 1$, então os polinômios $f_1, \ldots, f_n$ são ditos *coprimos*. Eles vão ser ditos *coprimos dois a dois* se $mdc(f_i, f_j) = 1$ para $1 \leq i < j \leq n$.

Podemos calcular o máximo divisor comum de dois polinômios $f, g \in F[x]$ utilizando o *algoritmo de Euclides*. Vamos supor, sem perda de generalidade, $g \neq 0$ e que g não divide f. Então, vamos aplicar repetidamente o algoritmo da divisão da seguinte forma

$$\begin{aligned}
f &= q_1 g + r_1 && 0 \leq \deg(r_1) < \deg(g) \\
g &= q_2 r_1 + r_2 && 0 \leq \deg(r_2) < \deg(r_1) \\
r_1 &= q_3 r_2 + r_3 && 0 \leq \deg(r_3) < \deg(r_2) \\
&= && \vdots \\
r_{s-2} &= q_2 r_{s-1} + r_s && 0 \leq \deg(r_s) < \deg(r_{s-1}) \\
r_{s-1} &= q_{s+1} r_s\,.
\end{aligned}$$

O processo se encerra pois a sequência de graus $\deg(g) > \deg(r_1) > \deg(r_2) > \ldots$ é estritamente decrescente, e como o grau é um inteiro não negativo, deve atingir o grau $-\infty$ (resto zero) em um número finito de passos.

O **Máximo Divisor Comum** dos polinômios f e g, denotado por $\mathrm{mdc}(f, g)$, é o último resto não nulo, r_s, na sequência de divisões.

Exemplo 46. *O algoritmo de Euclides aplicado aos polinômios*

$$f(x) = 2x^6 + x^3 + x^2 + 2 \quad e \quad g(x) = x^4 + x^2 + 2x\,,$$

em $\mathbb{F}_3[x]$*, nos dá*

$$\begin{aligned} 2x^6 + x^3 + x^2 + 2 &= (2x^2)(x^4 + x^2 + 2x) + (x^4 + x^2 + 2) \\ x^4 + x^2 + 2x &= (1)(-x^4 + x^2 + 2) + (2x + 1) \\ x^4 + x^2 + 2 &= (2x^3 + 2x^2 + x + 1)(2x + 1) + 1 \\ 2x + 1 &= (2x + 1) \cdot 1 \end{aligned}$$

Portanto, $mdc(f, g) = 1$ *e* f, g *são coprimos.*

Definição 47. *Um polinômio* $p \in F[x]$ *não nulo é dito* irredutível sobre F *(ou irredutível em* $F[x]$*, ou primo em* $F[x]$*) se* p *tem grau positivo e* $p = bc$*, com* $b, c \in F[x]$*, implica em* b *ou* c *ser o polinômio constante.*

Lema 48. *Se um polinômio irredutível* p *em* $F[x]$ *divide um produto* $f_1 \cdots f_m$ *de polinômios em* $F[x]$*, então algum dos fatores* f_j *é divisível por* p*.*

Teorema 49 (Fatoração Única em $F[x]$)**.** *Seja* $f \in F[x]$ *um polinômio qualquer de grau positivo. Então, ele pode ser escrito de forma única, a menos da ordem dos fatores,*

$$f = a p_1^{e_1} \cdots p_k^{e_k}\,, \tag{2.2}$$

onde $a \in F$*,* $p_1, \ldots, p_k$ *são polinômios mônicos irredutíveis em* $F[x]$ *e* $e_1, \ldots, e_k$ *são inteiros positivos.*

Demonstração. Faremos a demonstração por indução sobre o grau de f. O caso em que $\deg(f) = 1$ é direto, pois todo polinômio em $F[x]$ de grau 1 é irredutível em F. Agora, vamos supor que a fatoração (2.2) vale para todos os polinômios não constantes em $F[x]$ de grau menor que n. Se $\deg(f) = n$ e f é irredutível sobre F o resultado segue, pois podemos escrever $f = a(a^{-1}f)$, onde a é o coeficiente líder de f e $a^{-1}f$ é um polinômio mônico irredutível em $F[x]$. Caso contrário, f admite uma fatoração $f = gh$, com $1 \leq \deg(g) < n$, $1 \leq \deg(h) < n$ e $g, h \in F[x]$. Pela hipótese de indução, g e h tem fatoração como em (2.2) e, assim, f pode ser escrito dessa forma.

Agora, provaremos a unicidade da fatoração. Suponhamos que f tenha duas fatorações da forma (2.2)

$$f = a p_1^{e_1} \cdots p_k^{e_k} = b q_1^{d_1} \cdots q_r^{d_r}\,. \tag{2.3}$$

Comparando os coeficientes líderes, chegamos que $a = b$. Além disso, o polinômio irredutível p_1 em $F[x]$ divide o lado direito da equação (2.3). Daí, pelo Lema 48, existe j, com $1 \leq j \leq r$, tal que p_1 divide q_j. Porém, q_j é também irredutível em $F[x]$ e devemos

ter $q_j = cp_1$, onde c é um polinômio constante. Como p_1 e q_j são ambos polinômios mônicos, temos $q_j = p_1$ e assim eles se cancelam na equação (2.3). Podemos repetir o argumento no resto da equação e, após um número finito de passos, chegamos que as duas fatorações são iguais, a menos da ordem dos fatores. □

Como os polinômios irredutíveis sobre um corpo F são exatamente os elementos primos em $F[x]$, o resultado a seguir, que já foi em partes aplicado no Lema 48, é uma consequência imediata do item 4 do Teorema 38 e do Teorema 44.

Teorema 50. *Seja $f \in F[x]$. O anel quociente $F[x]/\langle f \rangle$ é um corpo se e somente se f é irredutível sobre F.*

Definição 51. *Um elemento $b \in F$ é dito uma* raiz *(ou um* zero*) do polinômio $f \in F[x]$ se $f(b) = 0$.*

Teorema 52. *Um elemento $b \in F$ é uma raiz do polinômio $f \in F[x]$ se e somente se $x - b$ divide $f(x)$.*

Definição 53. *Sejam $b \in F$ uma raiz de um polinômio $f \in F[x]$ e k um inteiro positivo. Se k é tal que $f(x)$ é divisível por $(x-b)^k$, mas não por $(x-b)^{k+1}$, então dizemos que k é a multiplicidade de b. Além disso, se $k = 1$, então b é dita uma* raiz simples *de f; se $k \geq 2$, então b é* raiz múltipla *de f.*

Teorema 54. *Seja $f \in F[x]$ com $\deg(f) = n \geq 0$. Se $b_1, \ldots, b_m \in F$ são raízes distintas de f com multiplicidades $k_1, \ldots, k_m$, respectivamente, então $(x - b_1)^{k_1} \cdots (x - b_m)^{k_m}$ divide $f(x)$. Consequentemente, $k_1 + \cdots + k_m \leq n$, e f pode ter no máximo n raízes distintas em F.*

Definição 55. *Se $f(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n \in F[x]$, então a* derivada *$f'$ de f em $F[x]$ é dada por $f' = f'(x) = a_1 + 2a_2 + \cdots + na_n x^{n-1}$.*

Teorema 56. *O elemento $b \in F$ é uma raiz múltipla do polinômio $f \in F[x]$ se e somente se é uma raiz de ambos f e sua derivada f'.*

Teorema 57. *O polinômio $f \in F[x]$ de grau 2 ou 3 é irredutível em $F[x]$ se e somente se f não tem raízes em F.*

Demonstração. Se f é um polinômio irredutível com grau 2 ou 3, então o Teorema 52 nos mostra que f não tem raízes em F. Por outro lado, se f não tem raízes em F e fosse redutível em $F[x]$, poderíamos escrever $f = gh$, com $g, h \in F[x]$ e $1 \leq \deg(g) \leq \deg(h)$. Porém, $\deg(g) + \deg(h) = \deg(f) \leq 3$. Logo, $\deg(g) = 1$, ou seja, $g(x) = ax + b$, com $a, b \in F$ e $a \neq 0$. Assim, $-ba^{-1}$ é uma raiz de g, mas também uma raiz de f em F, uma contradição. □

Definição 58. *Seja* $f \in R[x_1, \ldots, x_n]$ *dado por*

$$f(x_1, \ldots, x_n) = \sum a_{i_1 \cdots i_n} x_1^{i_1} \cdots x_n^{i_n} .$$

Se $a_{i_1 \cdots i_n} \neq 0$*, então* $a_{i_1 \ldots i_n} x_1^{i_1} \cdots x_n^{i_n}$ *é chamado um* termo *de* f *e* $i_1 + \cdots + i_n$ *é o grau do termo. Para* $f \neq 0$*, definimos o* grau *de* f*, denotado por* $\deg(f)$*, como o maior dos graus dos termos de* f*. Para* $f = 0$*, convencionamos* $\deg(f) = -\infty$*. Além disso, se* $f \neq 0$ *e todos os termos de* f *têm o mesmo grau, então dizemos que* f *é* homogêneo*.*

Definição 59. *Um polinômio* $f \in R[x_1, \ldots, x_n]$ *é dito* simétrico *se* $f(x_{i_1}, \ldots, x_{i_n}) = f(x_1, \ldots, x_n)$ *para toda permutação* $i_1, \ldots, i_n$ *dos inteiros* $1, \ldots, n$*.*

2.2 Extensão de Corpos

Seja F um corpo. Um subconjunto K de F é chamado de *subcorpo* de F quando é um corpo com as operações induzidas de F. Nesse contexto, F é dito uma *extensão de corpo* de K. Se $K \neq F$, dizemos que K é um *subcorpo próprio* de F.

Se K é um subcorpo do corpo finito $\mathbb{F}_p$, com p primo, então K contém os elementos 0 e 1 e daí todos os elementos de $\mathbb{F}_p$, pelo fato de K ser fechado sobre a adição. Logo, temos que $\mathbb{F}_p$ não tem subcorpos próprios, o que nos motiva à seguinte definição.

Definição 60. *Um corpo que não tem subcorpos próprios é chamado de* corpo primo*.*

Teorema 61. *O subcorpo primo de um corpo* F *é isomorfo ou ao corpo* $\mathbb{F}_p$ *ou a* $\mathbb{Q}$*, a depender de se a característica de* F *é prima ou zero.*

Definição 62. *Sejam* K *um subcorpo de um corpo* F *e* M *um subconjunto qualquer de* F*. Definimos o corpo* $K(M)$ *como a interseção de todos os subcorpos de* F *contendo ambos* K *e* M*. Chamamos* $K(M)$ *de extensão de corpo de* K *obtida pela* adjunção *dos elementos de* M*. Além disso, se o conjunto* $M = \{\theta_1, \ldots, \theta_n\}$ *é finito, escrevemos* $K(M) = K(\theta_1, \ldots, \theta_n)$*. Quando* M *consiste de um único elemento* $\theta \in F$ *então* $L = K(\theta)$ *é dita uma* extensão simples *de* K *e* θ *é dito* elemento primitivo *da extensão.*

Definição 63. *Sejam* K *um subcorpo de* F *e* $\theta \in F$*. Se* θ *é raiz de um polinômio não trivial com coeficientes em* K*, isto é, se*

$$a_n\theta^n + \cdots + a_1\theta + a_0 = 0,$$

com $a_i \in K$ *não todos nulos, então dizemos que* θ *é* algébrico *sobre* K*. Uma extensão* L *de* K *é dita* algébrica *sobre* K *(ou uma* extensão algébrica *de* K*) se todo elemento de* L *é algébrico sobre* K*.*

Definição 64. *Seja* $\theta \in F$ *um elemento algébrico sobre* K*. O polinômio mônico* $g \in K[x]$ *unicamente determinado que gera o ideal* $J = \{f \in K[x] \mid f(\theta) = 0\}$ *de* $K[x]$ *é chamado* polinômio minimal *de* θ *sobre* K*. Pelo grau de* θ *queremos dizer o grau de* g*.*

Teorema 65. *Se $\theta \in F$ é algébrico sobre K, então seu polinômio minimal g sobre K satisfaz as seguintes propriedades:*

1. *g é irredutível em $K[x]$;*
2. *dado $f \in K[x]$ temos $f(\theta) = 0$ se e somente se g divide f;*
3. *g é o polinômio mônico de menor grau em $K[x]$ que tem θ como raiz.*

Demonstração. Item 1. Na notação da Definição 64, temos $J \neq \langle 0 \rangle$, pois θ é algébrico sobre K. Então, pelo Teorema 44, existe um polinômio mônico $g \in K[x]$ unicamente determinado tal que J é igual ao ideal principal $\langle g \rangle$. Afirmamos que g é irredutível sobre $K[x]$. Com efeito, g tem grau positivo, pois tem θ como raiz. Além disso, se $g = h_1 h_2$ em $K[x]$, com $1 \leq \deg(h_i) < \deg(g)$ e $i = 1, 2$, então $0 = g(\theta) = h_1(\theta)h_2(\theta)$ implica que um dos polinômios h_1 ou h_2 são elementos de J e, assim, divisíveis por g, um absurdo.

Item 2. Segue da definição de g.

Item 3. Notemos o seguinte, qualquer polinômio mônico em $K[x]$ tendo θ como raiz é um múltiplo de g. Daí, ou tal polinômio é igual a g ou seu grau é maior que o grau de g. □

Definição 66. *Seja L uma extensão de corpo de K. Tomemos L como um espaço vetorial sobre K. Nessas condições se L é de dimensão finita, então L é dita uma* extensão finita *de K. A dimensão do espaço vetorial L sobre K é chamada de* grau *de L sobre K, denotada por $[L : K]$.*

Teorema 67. *Se L é uma extensão finita de K e M é uma extensão finita de L, então M é uma extensão finita de K tal que*

$$[M : K] = [M : L][L : K].$$

Demonstração. Sejam $[M : L] = m$, $[L : K] = n$. Além disso, tomemos $\{\alpha_1, \ldots, \alpha_m\}$ uma base de M sobre L e $\{\beta_1, \ldots, \beta_n\}$ uma base de L sobre K. Então, todo elemento $\alpha \in M$ é uma combinação linear

$$\alpha = \gamma_1\alpha_1 + \cdots + \gamma_m\alpha_m$$

com $\gamma_i \in L$, onde $1 \leq i \leq m$. Escrevendo cada γ_i em termos dos elementos β_j da base, obtemos

$$\alpha = \sum_{i=1}^{m} \gamma_i\alpha_i = \sum_{i=1}^{m} \left(\sum_{j=1}^{n} r_{ij}\beta_j \right) \alpha_i = \sum_{i=1}^{m} \sum_{j=1}^{n} r_{ij}\beta_j\alpha_i \,,$$

onde $r_{ij} \in K$. Precisamos somente mostrar que os mn elementos $\beta_j\alpha_i$, com $1 \leq i \leq m$ e $1 \leq j \leq n$, são linearmente independentes sobre K. Suponhamos que

$$\sum_{i=1}^{m} \sum_{j=1}^{n} s_{ij}\beta_j\alpha_i = 0$$

com os coeficientes $s_{ij} \in K$. Então

$$\sum_{i=1}^{m}\left(\sum_{j=1}^{n} s_{ij}\beta_j\right)\alpha_i = 0\,.$$

Como os elementos α_i são linearmente independentes sobre L, temos

$$\sum_{j=1}^{n} s_{ij}\beta_j = 0, \quad 1 \leq i \leq m.$$

Porém, como os β_j são linearmente independentes sobre K, concluímos que os coeficientes s_{ij} são todos zero. □

Teorema 68. *Toda extensão de corpo finita de K é algébrica sobre K.*

Demonstração. Seja L uma extensão finita de K e tomemos $[L : K] = m$. Para $\theta \in L$, os $m + 1$ elementos $1, \theta, \ldots, \theta^m$ são necessariamente linearmente dependentes sobre K. Logo, temos a relação

$$a_0 + a_1\theta + \cdots + a_m\theta^m = 0,$$

com os coeficientes $a_i \in K$ não todos nulos. Logo, temos que θ é algébrico sobre K. □

Teorema 69. *Sejam $\theta \in F$ algébrico de grau n sobre K e g o polinômio minimal de θ sobre K. Então:*

1. *$K(\theta)$ é isomorfo a $K[x]/\langle g\rangle$;*
2. *$[K(\theta) : K] = n$ e $\{1, \theta, \ldots, \theta^{n-1}\}$ é uma base de $K(\theta)$ sobre K;*
3. *todo elemento $\alpha \in K(\theta)$ é algébrico sobre K e seu grau sobre K é um divisor de n.*

Demonstração. Item 1. Consideremos a função $\tau\colon K[x] \to K(\theta)$ dada por $\tau(f) = f(\theta)$. Vemos que essa função é um homomorfismo de anéis e $\ker\tau = \{f \in K[x] \mid f(\theta) = 0\} = \langle g\rangle$, pela definição de polinômio minimal. Seja S a imagem de τ, isto é, S é o conjunto das expressões polinomiais em θ com coeficientes em K. Pelo Teorema do homomorfismo para anéis, temos que S é isomorfo ao quociente $K[x]/\langle g\rangle$. Mas, pelo item 1 do Teorema 65 e pelo Teorema 50, $K[x]/\langle g\rangle$ é um corpo e, assim, S também o é. Como $K \subseteq S \subseteq K(\theta)$ e $\theta \in S$, segue da definição de $K(\theta)$ que $S = K(\theta)$, como queríamos.

Item 2. Como $S = K(\theta)$, qualquer $\alpha \in K(\theta)$ pode ser escrito na forma $\alpha = f(\theta)$ para algum $f \in K[x]$. Pelo algoritmo da divisão, $f = qg + r$ com $q, r \in K[x]$ e $\deg(r) < \deg(g) = n$. Assim, $\alpha = f(\theta) = q(\theta)g(\theta) + r(\theta) = r(\theta)$ e α é uma combinação linear de $1, \theta, \ldots, \theta^{n-1}$ com coeficientes em K. Por outro lado, se $a_0 + a_1\theta + \cdots + a_{n-1}\theta^{n-1} = 0$, para certos $a_i \in K$, então o polinômio $h(x) = a_0 + a_1x + \cdots + a_{n-1}x^{n-1} \in K[x]$ tem θ como raiz. Assim, pelo item 2 do Teorema 65, h é um múltiplo de g. Além disso $\deg(h) < n$,

mas isso é somente possível se $h = 0$, ou seja, se todos os coeficientes $a_i = 0$. Portanto, os elementos $1, \theta, \cdots, \theta^{n-1}$ são linearmente independentes em K.

Item 3. O corpo $K(\theta)$ é uma extensão de corpo finita de K pelo item 2, daí $\alpha \in K(\theta)$ é algébrico sobre K, pelo Teorema 68. Além disso, $K(\alpha)$ é um subcorpo de $K(\theta)$. Se d é o grau de α sobre K, então, pelo item 2 e pelo Teorema 67, temos

$$n = [K(\theta) : K] = [K(\theta) : K(\alpha)][K(\alpha) : K] = [K(\theta) : K(\alpha)]d.$$

Portanto, d divide n. □

Teorema 70. *Seja $f \in K[x]$ irredutível sobre o corpo K. Então existe uma extensão algébrica simples de K com uma raiz de f sendo o elemento primitivo.*

Demonstração. Consideremos o anel quociente $L = K[x]/\langle f \rangle$, que, pelo Teorema 50, é um corpo. Os elementos em L são as classes de congruência $[h] = h + \langle f \rangle$, com $h \in K[x]$. Para todo $a \in K$ podemos considerar a classe de congruência $[a]$ dada pelo polinômio constante a e, se $a, b \in K$ são distintos, então $[a] \neq [b]$, visto que f é de grau positivo. Vejamos que a aplicação levando $a \mapsto [a]$ nos dá um isomorfismo entre K sobre um subcorpo $K' \subset L$ de tal forma que possamos identificar K' com K. Ou seja, podemos ver L como uma extensão de K. Além disso, para todo polinômio

$$h(x) = a_0 + a_1 x + \cdots + a_m x^m$$

em $K[x]$, temos

$$\begin{aligned}
[h] &= [a_0 + a_1 x + \cdots + a_m x^m] \\
&= [a_0] + [a_1][x] + \cdots + [a_m][x]^m \\
&= a_0 + a_1[x] + \cdots + a_m[x]^m \,,
\end{aligned}$$

pelas operações com classes de congruência e tomando $[a_i] = a_i$. Logo, qualquer elemento em L pode ser escrito como uma expressão polinomial em $[x]$ com coeficientes em K. Ora, qualquer corpo contendo ambos K e $[x]$ deve conter essas expressões polinomiais, segue que L é uma extensão simples de K obtida pela adjunção de $[x]$. Se $f = b_0 + b_1 x + \cdots + b_n x^n$, então

$$\begin{aligned}
f([x]) &= b_0 + b_1[x] + \cdots + b_n[x]^n \\
&= [b_0 + b_1 x + \cdots + b_n x^n] \\
&= [f] = [0] \,,
\end{aligned}$$

e $[x]$ é uma raiz de f e L é uma extensão algébrica simples de K. □

Teorema 71. *Sejam α e β duas raízes de um polinômio $f \in K[x]$ irredutível sobre K. Então existe um isomorfismo entre $K(\alpha)$ e $K(\beta)$ que leva α em β e mantém os outros elementos de K fixos.*

Definição 72. *Seja $f \in K[x]$ de grau positivo e F uma extensão de corpo de K. Então, diz-se que f se decompõe em fatores lineares em F se f pode ser escrito como um produto de fatores lineares em $F[x]$. Isto é, se existem elementos $\alpha_1, \alpha_2, \ldots, \alpha_n \in F$ tais que*

$$f(x) = a(x - \alpha_1)(x - \alpha_2) \cdots (x - \alpha_n),$$

onde a é o coeficiente líder de f. O corpo F é um corpo de decomposição de f sobre K se f se decompõe em fatores lineares em F e, se, além disso, $F = K(\alpha_1, \alpha_2, \ldots, \alpha_n)$.

Observe que um corpo de decomposição F de f sobre K é o menor corpo que contém todas as raízes de f: nenhum subcorpo próprio de F que seja uma extensão de K contém todas as raízes de f.

Teorema 73 (Existência e Unicidade do Corpo de Decomposição)**.** *Se K é um corpo e f é um polinômio qualquer de grau positivo em $K[x]$, então existe um corpo de decomposição de f sobre K. Quaisquer dois corpos de decomposição de f sobre K são isomorfos sob um isomorfismo que mantém os elementos de K fixos e permuta as raízes de f.*

3 Corpos finitos

3.1 Caracterização dos corpos finitos

Nos capítulos anteriores, introduzimos os corpos finitos a partir de exemplos fundamentais, como o corpo $\mathbb{F}_p$, obtido por meio do anel dos inteiros módulo um primo. Também discutimos a construção de extensões por adjunção de raízes de polinômios irredutíveis, permitindo formar corpos com mais elementos.

Neste capítulo, aprofundaremos a caracterização dos corpos finitos, investigando quais ordens são possíveis, como essas estruturas podem ser construídas e em que sentido são únicas a menos de isomorfismos. Além disso, analisaremos propriedades fundamentais desses corpos, com destaque para a estrutura do grupo multiplicativo e para o papel desempenhado pelos subcorpos. Essas considerações são centrais para compreender a organização e a aplicação dos corpos finitos na álgebra e em outras áreas da matemática.

Lema 74. *Seja F um corpo finito contendo um subcorpo K com q elementos. Então F possui q^m elementos, onde $m = [F : K]$.*

Demonstração. O corpo F é um espaço vetorial sobre K e, como F é finito, a dimensão m é finita. Se $[F : K] = m$, então existe uma base $\{b_1, b_2, \ldots, b_m\}$ de F sobre K. Cada elemento de F pode ser escrito de forma única como uma combinação linear:

$$a_1b_1 + a_2b_2 + \cdots + a_mb_m \quad \text{com } a_i \in K.$$

Como cada a_i pode assumir q valores distintos, F tem exatamente q^m elementos. □

Teorema 75. *Seja F um corpo finito. Então F possui p^n elementos, onde p é a característica de F e $n = [F : \mathbb{F}_p]$.*

Demonstração. Como F é finito, sua característica é um primo p pelo Corolário 36. Sabemos que existe um subcorpo primo de F que é isomorfo a $\mathbb{F}_p$, logo F é uma extensão finita de $\mathbb{F}_p$ com $n = [F : \mathbb{F}_p]$, logo pelo Lema 74 temos que F possui p^n elementos □

Como visto até o momento, dado um corpo primo $\mathbb{F}_p$ e um polinômio $f \in \mathbb{F}_p[x]$ irredutível de grau n, a adjunção de uma raiz de f a $\mathbb{F}_p$ produz um corpo finito com p^n elementos.

Entretanto, surge uma questão fundamental: para todo inteiro positivo n, existe sempre um polinômio irredutível de grau n em $\mathbb{F}_p[x]$? A resposta afirmativa a essa pergunta é essencial para garantir a existência de um corpo finito com p^n elementos, para qualquer

primo p e qualquer $n \in \mathbb{N}$. Para demonstrar esse resultado, adotaremos uma abordagem baseada nos resultados que se seguem.

Lema 76. *Se F é um corpo finito com q elementos, então todo $a \in F$ satisfaz $a^q = a$.*

Demonstração. A identidade é trivial para $a = 0$. Para $a \neq 0$, temos que, como o grupo multiplicativo F^* tem ordem $q-1$, então $a^{q-1} = 1$ e multiplicando ambos os lados por a, obtemos $a^q = a$. □

Lema 77. *Se F é um corpo finito com q elementos e $K \subseteq F$, então o polinômio $x^q - x \in K[x]$ se fatora em $F[x]$ como*

$$x^q - x = \prod_{a \in F} (x - a),$$

e F é o corpo de decomposição de $x^q - x$ sobre K.

Demonstração. O polinômio $x^q - x$ tem grau q e, pelo Lema 76, cada elemento de F é raiz dele. Como há q raízes distintas, ele se fatora completamente em $F[x]$, e F é o menor corpo onde isso ocorre. □

Teorema 78 (Existência e Unicidade dos Corpos Finitos)**.** *Para todo primo p e todo inteiro positivo n, existe um corpo finito com p^n elementos. Todo corpo finito com $q = p^n$ elementos é isomorfo ao corpo de decomposição de $x^q - x$ sobre $\mathbb{F}_p$.*

Demonstração. **Existência.**
Seja $q = p^n$. Considere o polinômio $f(x) = x^q - x \in \mathbb{F}_p[x]$. Seja F o corpo de decomposição de f sobre $\mathbb{F}_p$.

A derivada de f é $f'(x) = qx^{q-1} - 1 = -1$, pois $q = p^n \equiv 0 \mod p$. Como f' é um polinômio constante não possui raízes, logo pelo Teorema 56, f tem q raízes distintas em F.

Defina $S = \{\alpha \in F \mid \alpha^q - \alpha = 0\}$ conjunto das raízes de f. Podemos ver que $S \subset F$. Afirmamos que S é um corpo.

De fato,

1. $0, 1 \in S$, pois $0^q = 0$ e $1^q = 1$.
2. Se $\alpha, \beta \in S$, então $(\alpha - \beta)^q = \alpha^q - \beta^q = \alpha - \beta$ (usando o Teorema 37), logo $\alpha - \beta \in S$.
3. Se $\alpha, \beta \in S$ com $\beta \neq 0$, então $(\alpha\beta^{-1})^q = \alpha^q(\beta^q)^{-1} = \alpha\beta^{-1}$, logo $\alpha\beta^{-1} \in S$.

Como f se decompõe completamente em S temos que é corpo de decomposição, porém pela Definição 72 temos que F é o menor corpo que decompõe f, logo $F = S$. Assim, F é um corpo finito com q elementos.

Unicidade.
Seja F um corpo finito com $q = p^n$ elementos. Pelo Teorema 75, F tem característica p e contém $\mathbb{F}_p$ como subcorpo. Pelo Lema 77, F é o corpo de decomposição de $x^q - x$ sobre $\mathbb{F}_p$.

Pelo Teorema 73, quaisquer dois corpos de decomposição de $x^q - x$ sobre $\mathbb{F}_p$ são isomorfos. Portanto, F é isomorfo ao corpo construído anteriormente a partir de S. □

Teorema 79 (Critério para Subcorpos)**.** *Seja $\mathbb{F}_q$ o corpo finito com $q = p^n$ elementos. Então:*

- *Todo subcorpo de $\mathbb{F}_q$ tem ordem p^m, onde m é um divisor positivo de n.*
- *Reciprocamente, para cada divisor m de n, existe exatamente um subcorpo de $\mathbb{F}_q$ com p^m elementos.*

Demonstração. Seja K um subcorpo de $\mathbb{F}_q$. Pelo Lema 74, como $\mathbb{F}_q$ é um espaço vetorial de dimensão finita sobre K, temos que $q = |K|^d$ onde $d = [\mathbb{F}_q : K]$. Como $q = p^n$ e $|K|$ deve ser uma potência de p, pelo Teorema 75, digamos $|K| = p^m$, segue que:

$$p^n = (p^m)^d = p^{md}$$

o que implica $n = md$. Portanto, m divide n.

Reciprocamente, como $m \mid n$, temos $p^m - 1 \mid p^n - 1$,pois $\alpha - 1 \mid \alpha^d - 1$. Logo $x^{p^m-1} - 1 \mid x^{p^n-1} - 1$. Multiplicando ambos os lados por x temos:

$$x^{p^m} - x \mid x^{p^n} - x \quad \text{em } \mathbb{F}_p[x],$$

Seja agora $\mathbb{F}_q$ o corpo de decomposição de $x^{p^n} - x$ sobre $\mathbb{F}_p$ pelo Teorema 78. Como $x^{p^m} - x$ divide $x^{p^n} - x$, todas as raízes de $x^{p^m} - x$ estão contidas em $\mathbb{F}_q$. O conjunto dessas raízes forma exatamente o corpo $\mathbb{F}_{p^m}$, que é portanto um subcorpo de $\mathbb{F}_q$.

Para mostrar a unicidade, suponha que existam dois subcorpos distintos E e E' de $\mathbb{F}_q$, ambos com p^m elementos. Cada um desses subcorpos consiste precisamente das raízes de $x^{p^m} - x$ em $\mathbb{F}_q$. Como este polinômio tem no máximo p^m raízes distintas, a existência de dois subcorpos distintos implicaria que existem mais de p^m raízes em $\mathbb{F}_q$, o que é uma contradição. Portanto, o subcorpo com p^m elementos é único. □

A demonstração do Teorema 79 mostra que o único subcorpo de $\mathbb{F}_{p^n}$ com p^m elementos, onde m é um divisor positivo de n, é constituído exatamente pelas raízes do polinômio $x^{p^m} - x \in \mathbb{F}_p[x]$ em $\mathbb{F}_{p^n}$.

Teorema 80. *Para todo corpo finito $\mathbb{F}_q$, o grupo multiplicativo $\mathbb{F}_q^*$ dos elementos não nulos de $\mathbb{F}_q$ é cíclico.*

Demonstração. Assumimos $q \geq 3$, os casos $q = 1, 2$ são triviais. Seja $h = q - 1$ a ordem do grupo $\mathbb{F}_q^*$ e considere a fatoração prima:

$$h = p_1^{r_1} p_2^{r_2} \cdots p_m^{r_m}$$

Para cada fator primo p_i, o polinômio $x^{h/p_i} - 1$ tem no máximo h/p_i raízes em $\mathbb{F}_q$. Como $h/p_i < h$, existe algum elemento $a_i \in \mathbb{F}_q^*$ que não é raiz deste polinômio. Defina:

$$b_i = a_i^{h/p_i^{r_i}}$$

- Temos que, $b_i^{p_i^{r_i}} = a_i^h = 1$, logo pelo Teorema 10 a ordem de b_i divide $p_i^{r_i}$.
- Mas $b_i^{p_i^{r_i-1}} = a_i^{h/p_i} \neq 1$ pela escolha de a_i, portanto a ordem de b_i é exatamente $p_i^{r_i}$.

Defina $b = b_1 b_2 \cdots b_m$. Vamos mostrar que b tem ordem h.

Suponha por contradição que a ordem de b seja um divisor próprio de h. Então existe algum p_i tal que a ordem de b divide h/p_i. Sem perda de generalidade, assuma $i = 1$.

Temos:

$$1 = b^{h/p_1} = b_1^{h/p_1} b_2^{h/p_1} \cdots b_m^{h/p_1}$$

Para $j \geq 2$, observe que $p_j^{r_j}$ divide h/p_1, pois $p_1 \neq p_j$, logo:

$$b_j^{h/p_1} = (b_j^{p_j^{r_j}})^{h/(p_1 p_j^{r_j})} = 1^{h/(p_1 p_j^{r_j})} = 1$$

Portanto, a equação se reduz a:

$$1 = b_1^{h/p_1}$$

Mas isto implicaria que a ordem de b_1, que é igual a $p_1^{r_1}$, divide h/p_1, o que é uma contradição pois $p_1^{r_1}$ não divide $h/p_1 = p_1^{r_1-1} p_2^{r_2} \cdots p_m^{r_m}$.

Logo o elemento b tem ordem $h = q - 1$ e portanto gera $\mathbb{F}_q^*$, que é assim um grupo cíclico. □

Definição 81. *Um elemento α de um corpo finito F é chamado de **primitivo** se ele gera o grupo multiplicativo F^*, isto é, se todo elemento não nulo de F pode ser escrito como uma potência de α.*

Teorema 82. *Seja $\mathbb{F}_q$ um corpo finito e $\mathbb{F}_r$ uma extensão finita de $\mathbb{F}_q$. Então $\mathbb{F}_r$ é uma extensão algébrica simples de $\mathbb{F}_q$ e todo elemento primitivo de $\mathbb{F}_r$ pode servir como elemento gerador de $\mathbb{F}_r$ sobre $\mathbb{F}_q$.*

Demonstração. Seja ζ um elemento primitivo de $\mathbb{F}_r$. Temos claramente que $\mathbb{F}_q(\zeta) \subset \mathbb{F}_r$. Por outro lado, como $\mathbb{F}_q(\zeta)$ contém 0 e todas as potências de ζ, que pelo Teorema 80 $\langle\zeta\rangle = \mathbb{F}_r^*$, e portanto $\mathbb{F}_r \subset \mathbb{F}_q(\zeta)$. Logo $\mathbb{F}_r = \mathbb{F}_q(\zeta)$.

Se η é outro elemento primitivo de $\mathbb{F}_r$, então η gera $\mathbb{F}_r^*$ e consequentemente $\mathbb{F}_r = \mathbb{F}_q(\eta)$. □

Corolário 83. *Para todo corpo finito $\mathbb{F}_q$ e todo inteiro positivo n, existe um polinômio irredutível em $\mathbb{F}_q[x]$ de grau n.*

Demonstração. Seja $\mathbb{F}_r$ uma extensão do corpo $\mathbb{F}_q$ com q^n elementos, de modo que $[\mathbb{F}_r : \mathbb{F}_q] = n$. Pelo Teorema 82, temos $\mathbb{F}_r = \mathbb{F}_q(\zeta)$ para algum $\zeta \in \mathbb{F}_r$. O polinômio minimal de ζ sobre $\mathbb{F}_q$ é irredutível em $\mathbb{F}_q[x]$ e tem grau n pelos Teoremas 65(1) e 69(2). □

3.2 Raízes de polinômios irredutíveis

Polinômios irredutíveis não podem ser fatorados em outros polinômios de grau menor e com coeficientes do mesmo corpo. Nesta seção exploramos este fato em relação as extensões de corpos.

Lema 84. *Seja $f \in \mathbb{F}_q[x]$ um polinômio irredutível sobre $\mathbb{F}_q$ e seja α uma raiz de f em uma extensão de $\mathbb{F}_q$. Então, para um polinômio $h \in \mathbb{F}_q[x]$, temos $h(\alpha) = 0$ se e somente se f divide h.*

Demonstração. Seja a o coeficiente líder de f e considere $g(x) = a^{-1}f(x)$. Então g é um polinômio mônico irredutível em $\mathbb{F}_q[x]$ com $g(\alpha) = 0$, e portanto é o polinômio minimal de α sobre $\mathbb{F}_q$ pela Definição 64. O resultado segue então do Teorema 65(ii). □

Lema 85. *Seja $f \in \mathbb{F}_q[x]$ um polinômio irredutível sobre $\mathbb{F}_q$ de grau m. Então $f(x)$ divide $x^{q^n} - x$ se e somente se m divide n.*

Demonstração. ($\Rightarrow$) Suponha que $f(x)$ divide $x^{q^n} - x$. Seja α uma raiz de f no corpo de decomposição de f sobre $\mathbb{F}_q$. Então $\alpha^{q^n} = \alpha$, logo $\alpha \in \mathbb{F}_{q^n}$. Segue que $\mathbb{F}_q(\alpha)$ é um subcorpo de $\mathbb{F}_{q^n}$. Como $[\mathbb{F}_q(\alpha) : \mathbb{F}_q] = m$ e $[\mathbb{F}_{q^n} : \mathbb{F}_q] = n$, o Teorema 67 mostra que m divide n.

($\Leftarrow$) Se m divide n, então o Teorema 79 implica que $\mathbb{F}_{q^n}$ contém $\mathbb{F}_{q^m}$ como subcorpo. Se α é uma raiz de f no corpo de decomposição de f sobre $\mathbb{F}_q$, então $[\mathbb{F}_q(\alpha) : \mathbb{F}_q] = m$, e portanto $\mathbb{F}_q(\alpha) = \mathbb{F}_{q^m}$. Consequentemente, temos $\alpha \in \mathbb{F}_{q^n}$, donde $\alpha^{q^n} = \alpha$, e assim α é raiz de $x^{q^n} - x \in \mathbb{F}_q[x]$. Pelo Lema 84, concluímos que $f(x)$ divide $x^{q^n} - x$.

□

Teorema 86. *Seja $f \in \mathbb{F}_q[x]$ um polinômio irredutível sobre $\mathbb{F}_q$ de grau m. Então:*

(i) f *tem uma raiz* α *em* $\mathbb{F}_{q^m}$.

(ii) Todas as raízes de f *são simples e são dadas pelos* m *elementos distintos* $\alpha, \alpha^q, \alpha^{q^2}, \ldots, \alpha^{q^{m-1}}$ *de* $\mathbb{F}_{q^m}$.

Demonstração. i) Seja α uma raiz de f no corpo de decomposição de f sobre $\mathbb{F}_q$. Então $[\mathbb{F}_q(\alpha) : \mathbb{F}_q] = m$, logo $\mathbb{F}_q(\alpha) = \mathbb{F}_{q^m}$ e em particular $\alpha \in \mathbb{F}_{q^m}$.

(ii) Primeiro mostramos que se $\beta \in \mathbb{F}_{q^m}$ é raiz de f, então β^q também é raiz de f. Escreva $f(x) = \sum_{i=0}^{m} a_i x^i$ com $a_i \in \mathbb{F}_q$. Pelo Lema 76 e Teorema 37 temos:

$$f(\beta^q) = \sum_{i=0}^{m} a_i \beta^{qi} = \sum_{i=0}^{m} a_i^q \beta^{qi} = \left(\sum_{i=0}^{m} a_i \beta^i \right)^q = f(\beta)^q = 0$$

Portanto, $\alpha, \alpha^q, \ldots, \alpha^{q^{m-1}}$ são raízes de f.

Agora provamos que estas raízes são distintas. Suponha por contradição que $\alpha^{q^j} = \alpha^{q^k}$ para $0 \leq j < k \leq m-1$. Elevando à potência q^{m-k}, obtemos:

$$\alpha^{q^{m-k+j}} = \alpha^{q^m} = \alpha$$

pelo Lema 84, $f(x)$ divide $x^{q^{m-k+j}} - x$. Pelo Lema 85, isto implica que m divide $m-k+j$, o que é impossível pois $0 < m-k+j < m$. □

Corolário 87. *Seja* $f \in \mathbb{F}_q[x]$ *um polinômio irredutível sobre* $\mathbb{F}_q$ *de grau* m. *Então o corpo de decomposição de* f *sobre* $\mathbb{F}_q$ *é* $\mathbb{F}_{q^m}$.

Demonstração. Pelo Teorema 86, f decompõe-se completamente em $\mathbb{F}_{q^m}$. Além disso, para qualquer raiz α de f em $\mathbb{F}_{q^m}$, temos que $\mathbb{F}_q(\alpha, \alpha^q, \ldots, \alpha^{q^{m-1}}) = \mathbb{F}_q(\alpha) = \mathbb{F}_{q^m}$, onde a última igualdade segue da demonstração do Teorema 86. Portanto, $\mathbb{F}_{q^m}$ é o corpo de decomposição de f sobre $\mathbb{F}_q$. □

Corolário 88. *Quaisquer dois polinômios irredutíveis em* $\mathbb{F}_q[x]$ *do mesmo grau possuem corpos de decomposição isomorfos.*

Definição 89. *Seja* $\mathbb{F}_{q^m}$ *uma extensão de* $\mathbb{F}_q$ *e seja* $\alpha \in \mathbb{F}_{q^m}$. *Então os elementos* $\alpha, \alpha^q, \alpha^{q^2}, \ldots, \alpha^{q^{m-1}}$ *são chamados de* conjugados *de* α *com respeito a* $\mathbb{F}_q$.

Os conjugados de $\alpha \in \mathbb{F}_{q^m}$ com respeito a $\mathbb{F}_q$ são distintos se e somente se o polinômio minimal de α sobre $\mathbb{F}_q$ tem grau m. Caso contrário, o grau d deste polinômio minimal é um divisor próprio de m, e os conjugados de α com respeito a $\mathbb{F}_q$ são os d elementos distintos $\alpha, \alpha^q, \ldots, \alpha^{q^{d-1}}$, cada um repetido m/d vezes.

Teorema 90. *Os conjugados de* $\alpha \in \mathbb{F}_q^*$ *com respeito a qualquer subcorpo de* $\mathbb{F}_q$ *têm a mesma ordem no grupo* $\mathbb{F}_q^*$.

Demonstração. Como $\mathbb{F}_q^*$ é um grupo cíclico pelo Teorema 80, o resultado segue diretamente do Teorema 11(ii), observando que qualquer potência da característica de $\mathbb{F}_q$ é coprima com a ordem $q-1$ de $\mathbb{F}_q^*$.

□

Corolário 91. *Se α é um elemento primitivo de $\mathbb{F}_q$, então todos os seus conjugados com respeito a qualquer subcorpo de $\mathbb{F}_q$ também são elementos primitivos de $\mathbb{F}_q$.*

Exemplo 92. *Seja $\alpha \in \mathbb{F}_{16}$ uma raiz de $f(x) = x^4 + x + 1 \in \mathbb{F}_2[x]$. Então os conjugados de α com respeito a $\mathbb{F}_2$ são*

$$\alpha,\ \alpha^2,\ \alpha^4 = \alpha + 1,\ \alpha^8 = \alpha^2 + 1,$$

a ultima igualdade vale pelo Teorema 37, e temos que cada um deles sendo um elemento primitivo de $\mathbb{F}_{16}$. Os conjugados de α com respeito a $\mathbb{F}_4$ são α e $\alpha^4 = \alpha + 1$

Existe uma relação íntima entre elementos conjugados e certos automorfismos de um corpo finito. Seja $\mathbb{F}_{q^m}$ uma extensão de $\mathbb{F}_q$. Por um automorfismo σ de $\mathbb{F}_{q^m}$ sobre $\mathbb{F}_q$, entende-se um automorfismo de $\mathbb{F}_{q^m}$ que fixa os elementos de $\mathbb{F}_q$. Em detalhes, exige-se que σ seja uma função bijetora de $\mathbb{F}_{q^m}$ em si mesma, tal que:

- $\sigma(a+b) = \sigma(a) + \sigma(b)$,
- $\sigma(ab) = \sigma(a)\sigma(b)$, para todos $a, b \in \mathbb{F}_{q^m}$,
- $\sigma(a) = a$, para todo $a \in \mathbb{F}_q$.

Teorema 93. *Os automorfismos distintos de $\mathbb{F}_{q^m}$ sobre $\mathbb{F}_q$ são exatamente as aplicações $\sigma_0, \sigma_1, \ldots, \sigma_{m-1}$, definidas por $\sigma_j(\alpha) = \alpha^{q^j}$ para $\alpha \in \mathbb{F}_{q^m}$ e $0 \leq j \leq m-1$.*

Demonstração. Para cada σ_j e $\alpha, \beta \in \mathbb{F}_{q^m}$:

1. $\sigma_j(\alpha\beta) = (\alpha\beta)^{q^j} = \alpha^{q^j}\beta^{q^j} = \sigma_j(\alpha)\sigma_j(\beta)$
2. $\sigma_j(\alpha+\beta) = (\alpha+\beta)^{q^j} = \alpha^{q^j} + \beta^{q^j} = \sigma_j(\alpha) + \sigma_j(\beta)$, pelo Teorema 37
3. $\sigma_j(a) = a^{q^j} = a$ para todo $a \in \mathbb{F}_q$,pelo Lema 76

Além disso, σ_j é bijetora pois é injetora,$\sigma_j(\alpha) = 0 \Rightarrow \alpha = 0$, e $\mathbb{F}_{q^m}$ é finito.

Sejam $0 \leq j < k \leq m-1$. Tome ζ um elemento primitivo de $\mathbb{F}_{q^m}$. Como $\zeta^{q^j} \neq \zeta^{q^k}$ para $j \neq k$, segue que $\sigma_j \neq \sigma_k$.

Seja σ um automorfismo arbitrário de $\mathbb{F}_{q^m}$ sobre $\mathbb{F}_q$. Seja β um elemento primitivo de $\mathbb{F}_{q^m}$ e $f(x) = x^m + \sum_{i=0}^{m} a_i x^i \in \mathbb{F}_q[x]$ seu polinômio minimal. Então:

$$0 = \sigma(f(\beta)) = \sigma(\beta)^m + \sum_{i=0}^{m-1} a_i \sigma(\beta)^i = f(\sigma(\beta))$$

Logo, como $\sigma(\beta)$ é raiz de f e pelo Teorema 86, $\sigma(\beta) = \beta^{q^j}$ para algum $0 \leq j \leq m-1$. Como σ é completamente determinado por sua ação no gerador β, temos $\sigma = \sigma_j$. □

3.3 Raízes da unidade e polinômio ciclotômico

Nesta seção, estudamos as raízes da unidade em corpos finitos e sua relação com a estrutura multiplicativa desses corpos. Apresentamos os polinômios ciclotômicos e discutimos como suas raízes descrevem subgrupos cíclicos do grupo multiplicativo. Esses conceitos são fundamentais para a análise de extensões de corpos e para aplicações na teoria dos códigos e na aritmética computacional.

Definição 94. *Seja n um inteiro positivo. O corpo de decomposição de $x^n - 1$ sobre um corpo K é chamado de* n-ésimo corpo ciclotômico *sobre K e denotado por $K^{(n)}$. As raízes de $x^n - 1$ em $K^{(n)}$ são chamadas de* n-ésimas raízes da unidade *sobre K e o conjunto de todas essas raízes é denotado por $E^{(n)}$.*

Um caso especial dessa Definição geral ocorre quando K é o corpo dos números racionais. Nesse caso, $K^{(n)}$ é um subcorpo do corpo dos números complexos, e as raízes n-ésimas da unidade possuem sua conhecida interpretação geométrica como os vértices de um polígono regular com n lados, localizado na circunferência unitária do plano complexo.

Para nossos propósitos, o caso mais importante é quando K é um corpo finito. No entanto, as propriedades básicas das raízes da unidade podem ser estabelecidas sem essa restrição. A estrutura de $E^{(n)}$ é determinada pela relação entre n e a característica de K, como será mostrado no Teorema a seguir. Quando nos referimos à característica p de K nesta discussão, permitimos também o caso $p = 0$.

Teorema 95. *Seja n um inteiro positivo e K um corpo de característica p. Então:*

(i) Se p não divide n, então $E^{(n)}$ é um grupo cíclico de ordem n com respeito à multiplicação em $K^{(n)}$.

(ii) Se p divide n, escreva $n = mp^e$ com tanto m quanto e inteiros positivos onde p não divide m. Então $K^{(n)} = K^{(m)}$, $E^{(n)} = E^{(m)}$, e as raízes de $x^n - 1$ em $K^{(n)}$ são os m elementos de $E^{(m)}$, cada um com multiplicidade p^e.

Demonstração. **Caso (i):** Quando p não divide n

O polinômio $x^n - 1$ e sua derivada nx^{n-1} não têm raízes em comum, pois nx^{n-1} só tem a raiz 0 em $K^{(n)}$ e $0^n - 1 \neq 0$. Pelo Teorema 56, $x^n - 1$ não tem raízes múltiplas, logo $|E^{(n)}| = n$.

Para $\zeta, \eta \in E^{(n)}$, temos $(\zeta\eta^{-1})^n = \zeta^n(\eta^n)^{-1} = 1$, pois $\zeta^n - 1 = 0 \Rightarrow \zeta^n = 1$ e o mesmo vale para η, logo $\zeta\eta^{-1} \in E^{(n)}$, mostrando que $E^{(n)}$ é subgrupo multiplicativo.

Seja $n = \prod_{i=1}^{t} p_i^{e_i}$ a fatoração prima de n. Para cada $p_i^{e_i}$, existe $\alpha_i \in E^{(n)}$ que não é raiz de $x^{n/p_i} - 1$. Então $\beta_i = \alpha_i^{n/p_i^{e_i}}$ tem ordem $p_i^{e_i}$. O produto $\beta = \prod_{i=1}^{t} \beta_i$ tem ordem n e gera $E^{(n)}$.

Caso (ii): Quando p divide n

Escrevendo $n = mp^e$ com $p \nmid m$, temos:

$$x^n - 1 = x^{mp^e} - 1 = (x^m - 1)^{p^e}$$

As raízes de $x^n - 1$ coincidem com as de $x^m - 1$, mas com multiplicidade p^e. Portanto $E^{(n)} = E^{(m)}$ e $K^{(n)} = K^{(m)}$ como corpos de decomposição. □

Definição 96. *Seja K um corpo de característica p e n um inteiro positivo não divisível por p. Então um gerador do grupo cíclico $E^{(n)}$ é chamado de* raiz primitiva n-ésima da unidade *sobre K.*

Sob as hipóteses da Definição anterior, o Teorema 11(v) garante que existem exatamente $\varphi(n)$ raízes primitivas n-ésimas da unidade em K. Escolhendo uma raiz primitiva ζ, todas as demais são dadas pelas potências ζ^s, com $1 < s < n$ e $mdc(s, n) = 1$. O polinômio que possui exatamente essas $\varphi(n)$ raízes como zeros é de grande interesse teórico e prático.

Definição 97. *Seja K um corpo de característica p, n um inteiro positivo não divisível por p, e ζ uma raiz primitiva n-ésima da unidade sobre K. Então o polinômio*

$$Q_n(x) = \prod_{\substack{1 \leq s \leq n \\ mdc(s,n)=1}} (x - \zeta^s)$$

é chamado de n-ésimo polinômio ciclotômico *sobre K.*

O polinômio $Q_n(x)$, cuja Definição envolve uma raiz primitiva ζ, é invariante em relação à escolha dessa raiz. Ele possui grau $\varphi(n)$, e seus coeficientes pertencem ao corpo ciclotômico de ordem n sobre K. No entanto, é possível demonstrar facilmente que tais coeficientes estão, de fato, contidos no subcorpo primo de K. Adotamos a notação $\prod_{d|n}$ para denotar o produto tomado sobre todos os divisores positivos d de um inteiro positivo n.

Teorema 98. *Seja K um corpo de característica p e n um inteiro positivo não divisível por p. Então:*

(i) $x^n - 1 = \prod_{d|n} Q_d(x)$

(ii) Os coeficientes de $Q_n(x)$ pertencem ao subcorpo primo de K, e a $\mathbb{Z}$ se o subcorpo primo de K é o corpo dos números racionais.

Demonstração. (i)

Cada raiz n-ésima da unidade ζ^s, com uma ζ raiz primitiva, é uma raiz primitiva d-ésima para exatamente um divisor d de n, onde $d = n/mdc(s, n)$.

Agrupando os fatores $(x - \zeta^s)$ de acordo com o valor d correspondente:

$$x^n - 1 = \prod_{s=1}^{n}(x - \zeta^s) = \prod_{d|n} \prod_{\substack{1\leq s\leq n \\ mdc(s,n)=n/d}} (x - \zeta^s) = \prod_{d|n} Q_d(x)$$

(ii)

Aqui faremos uma indução sobre n. Para $n = 1$ temos:

$$Q_1(x) = x - 1$$

que satisfaz as condições.

Assuma válido para todo $k < n$.

$$x^n - 1 = Q_n(x) \cdot \prod_{\substack{d|n \\ d<n}} Q_d(x) = Q_n(x) \cdot f(x)$$

onde $f(x)$ contém todos os $Q_d(x)$ com $d|n$ e $d < n$.

Pela hipótese indutiva, cada $Q_d(x)$ com $d < n$ tem coeficientes no subcorpo primo (ou em $\mathbb{Z}$). Assim o produto $f(x)$ preserva esta propriedade

Como $f(x)$ é mônico (produto de polinômios mônicos). Podemos dividir $x^n - 1$ por $f(x)$ usando o algoritmo de divisão. Portanto os coeficientes do quociente $Q_n(x)$ permanecem no mesmo anel

Para o caso particular se K tem subcorpo primo $\mathbb{Q}$, os coeficientes são inteiros algébricos, porém como estão em $\mathbb{Q}$ e são inteiros algébricos, devem ser inteiros

□

Teorema 99. *O corpo ciclotômico $K^{(n)}$ é uma extensão algébrica simples de K. Além disso:*

(i) Se $K = \mathbb{Q}$, então o polinômio ciclotômico Q_n é irredutível sobre K e $[K^{(n)} : K] = \phi(n)$.

(ii) Se $K = \mathbb{F}_q$ com $mdc(q, n) = 1$, então Q_n fatora-se em $\phi(n)/d$ polinômios mônicos irredutíveis distintos em $K[x]$, todos de mesmo grau d, $K^{(n)}$ é o corpo de decomposição de qualquer um desses fatores irredutíveis sobre K, e $[K^{(n)} : K] = d$, onde d é o menor inteiro positivo tal que $q^d \equiv 1 \bmod n$.

Demonstração. Omitiremos a demonstração de (i), para não fugirmos do tema do trabalho, porém pode ser encontrada em (LIDL; NIEDERREITER, 1997) página 65.

(ii) Seja ζ uma raiz primitiva n-ésima da unidade sobre $\mathbb{F}_q$.

O polinômio minimal de ζ sobre $\mathbb{F}_q$ tem grau d igual ao menor inteiro para o qual $\zeta^{q^d} = \zeta$, o que ocorre se e somente se $q^d \equiv 1 \pmod{n}$. Como Q_n tem $\phi(n)$ raízes distintas, pelo Teorema 95 e cada fator irredutível tem d raízes conjugadas, pelo Teorema 86, o número de fatores é $\frac{\phi(n)}{d}$.

O corpo $K^{(n)} = \mathbb{F}_q(\zeta)$ contém todas as raízes de Q_n e é portanto o corpo de decomposição, com dimensão d sobre $\mathbb{F}_q$ pelo Teorema 82. □

Exemplo 100. *Considere $K = \mathbb{F}_{11}$ e o polinômio ciclotômico $Q_{12}(x) = x^4 - x^2 + 1 \in \mathbb{F}_{11}[x]$. Na notação do Teorema anterior, temos $d = 2$. Em detalhes, $Q_{12}(x)$ se fatora como:*

$$Q_{12}(x) = (x^2 + 5x + 1)(x^2 - 5x + 1)$$

onde ambos os fatores são irredutíveis em $\mathbb{F}_{11}[x]$. O corpo ciclotômico $K^{(12)}$ é igual a $\mathbb{F}_{121}$.

Teorema 101. *O corpo finito $\mathbb{F}_q$ é o $(q-1)$-ésimo corpo ciclotômico sobre qualquer um de seus subcorpos.*

Demonstração. O polinômio $x^{q-1} - 1$ se decompõe completamente em $\mathbb{F}_q$ pois suas raízes são exatamente os elementos não nulos de $\mathbb{F}_q$, que formam o grupo multiplicativo $\mathbb{F}_q^*$ de ordem $q - 1$.

Temos que $\mathbb{F}_q$ é o menor corpo contendo todas as raízes de $x^{q-1} - 1$, pois qualquer subcorpo próprio de $\mathbb{F}_q$ teria menos elementos e não conteria todas as raízes.

Pela Definição 94, o corpo de decomposição de $x^{q-1} - 1$ sobre qualquer subcorpo $\mathbb{F}_{q_0}$ de $\mathbb{F}_q$ é exatamente $\mathbb{F}_q$, que é portanto o $(q-1)$-ésimo corpo ciclotômico sobre $\mathbb{F}_{q_0}$. □

Pelo Teorema 80, o grupo multiplicativo F^* possui estrutura cíclica de ordem $q - 1$. Assim, para cada divisor positivo n de $q-1$, existe em F^* um subgrupo cíclico de ordem n, cujos elementos podem ser descritos como $\{1, \alpha, \alpha^2, \ldots, \alpha^{n-1}\}$, conforme o Teorema 11(iii).

Esses elementos são exatamente as raízes n-ésimas da unidade sobre qualquer subcorpo de $\mathbb{F}_q$, sendo α, o gerador do subgrupo, uma raiz primitiva n-ésima da unidade sobre esse mesmo subcorpo.

Lema 102. *Se d é um divisor do inteiro positivo n com $1 \leq d < n$, então $Q_n(x)$ divide $(x^n - 1)/(x^d - 1)$ sempre que $Q_n(x)$ estiver definido.*

Demonstração. Pelo Teorema 98(i), temos a fatoração:

$$x^n - 1 = Q_n(x) \cdot \prod_{\substack{k|n \\ k \neq n}} Q_k(x)$$

$$x^d - 1 = \prod_{k|d} Q_k(x)$$

Como d é divisor próprio de n, todo divisor k de d também é divisor próprio de n. Portanto, cada $Q_k(x)$ no denominador aparece no numerador.

Assim podemos escrever:

$$\frac{x^n - 1}{x^d - 1} = Q_n(x) \cdot \prod_{\substack{k|n \\ k \nmid d}} Q_k(x)$$

o que mostra explicitamente que $Q_n(x)$ é fator do quociente. □

4 Algumas aplicações

Neste capítulo, o foco principal será a aplicação dos corpos finitos em criptografia, explorando como suas propriedades algébricas são utilizadas na construção de sistemas criptográficos modernos. Serão discutidos alguns exemplos e métodos que ilustram o uso desses corpos na proteção e codificação de informações. Além disso, será apresentada uma introdução à teoria dos códigos, abordada de forma breve e sem aprofundamento teórico, com o intuito de exemplificar como os mesmos conceitos algébricos também se aplicam à correção e detecção de erros em comunicações.

A principal referência para esta parte foi: (LIDL; PILZ, 1997) e (PANARIO, 2007).

4.1 Criptografia

Nesta seção, são apresentados de forma introdutória os princípios matemáticos que sustentam os criptossistemas, com foco nas aplicações de corpos finitos

4.1.1 Introdução à criptografia

A criptografia consiste em transformar uma mensagem original (*plaintext*) em uma forma codificada (*ciphertext*) por meio de uma função de encriptação, de modo que apenas o destinatário autorizado possa recuperá-la. O processo depende de um algoritmo e de uma chave conhecida apenas pelo emissor e pelo receptor. A operação inversa é chamada *decifragem*, realizada pela função de decodificação.

Nos sistemas clássicos ou *simétricos*, utiliza-se a mesma chave para cifrar e decifrar, enquanto os sistemas *assimétricos* ou de *chave pública* empregam chaves distintas. A segurança de um criptossistema baseia-se na dificuldade computacional de descobrir a chave ou o texto original sem autorização, mesmo conhecendo o método de encriptação.

A Figura (1) mostra o diagrama da criptografia clássica.

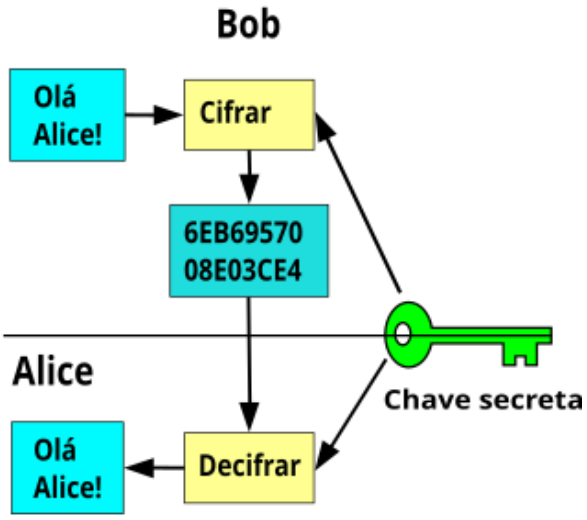

Figura 1 – Diagrama de chave simétrica. Acessado em https://pt.wikipedia.org/wiki/Criptografia.

Na criptografia clássica, uma das transformações básicas da mensagem original é a *cifras de substituição*, que trocam cada letra do texto por outra do alfabeto cifrado. Se cada letra do texto original corresponder à uma única letra, o sistema é denominado *monoalfabético*. Sistemas em que uma mesma letra do texto original pode corresponder a mais de uma letra são chamados *polialfabéticos*; esses últimos usam uma sequência de ciframentos monoalfabéticos.

Para representar as mensagens numericamente, costuma-se identificar letras com inteiros (por exemplo, $A = 0, \ldots, Z = 25$) ou agrupar letras em blocos , representando um par NO por $N \cdot 26 + O = 13 \cdot 26 + 14 = 252$.

Seja $\mathcal{A}$ uma sequência de caracteres do texto original e $\mathcal{B}$ a sequência de caracteres do texto cifrado; uma *chave* é uma aplicação injetiva $f : \mathcal{A} \longrightarrow \mathcal{B}$ definida como $a_1 a_2 \ \ldots \ a_n \longmapsto f(a_1) f(a_2) \ \ldots \ f(a_n)$. Se a mesma função é usada para todas as letras, tem-se uma *chave fixa*; caso varie conforme parâmetros como o tempo ou o próprio texto, trata-se de uma *chave variável*. Essas ideias permitem a construção de diversos métodos de encriptação baseados em funções lineares, matrizes, permutações, produtos escalares e no uso de números primos grandes.

Definição 103. *A aplicação $f : \mathbb{Z}_q \to \mathbb{Z}_q$, definida por*

$$a \mapsto na + k,$$

com $n, k \in \mathbb{Z}_q$ fixos e $mdc(n, q) = 1$, é chamada de cifra modular. O par (n, k) é a chave.

Para $n = 1$, a cifra é chamada Cifra de César; para $n > 1$ e $mdc(n, q) = 1$, a cifra resultante é denominada Cifra Afim.

Note que, no caso particular em que $n = 1$, a cifra modular reduz-se a uma translação das letras por k posições no alfabeto. Esse é justamente o princípio da Cifra de César, na qual cada letra é substituída pela que se encontra k posições adiante.

Observe também que, para decifrarmos um texto dado por

$$b = na + k \quad \text{em } \mathbb{Z}_q,$$

basta isolar a em $\mathbb{Z}_q$, obtendo:

$$a = (b - k)n^{-1} \quad \text{em } \mathbb{Z}_q.$$

Assim, a decifragem consiste em subtrair o valor k e multiplicar o resultado pelo inverso multiplicativo de n módulo q.

Exemplo 104. *Seja $\mathcal{A} = \mathbb{Z}_{26}$, e representemos as letras do alfabeto da seguinte forma:*

A	***B***	***C***	***D***	***E***	***F***	***G***	***H***	***I***	***J***	***K***	***L***	***M***	***N***	***O***	***P***	***Q***	***R***	***S***	***T***	***U***	***V***	***W***	***X***	***Y***	***Z***
0	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*	*13*	*14*	*15*	*16*	*17*	*18*	*19*	*20*	*21*	*22*	*23*	*24*	*25*

Para a cifra modular $a \mapsto 5a + 4$ em $\mathbb{Z}_{26}$, a palavra ***BUSSOLA****, isto é,*

$$(1, 20, 18, 18, 14, 11, 0),$$

é cifrada como:

$$b = 5a + 4 \pmod{26}.$$

Assim, obtemos $(9, 0, 16, 16, 22, 7, 4)$, o que corresponde, em letras, à palavra ***JAQQWHE****.*

Agora, considerando o texto cifrado ***NAMJYLUM****, representado por*

$$(13, 0, 12, 9, 24, 11, 21, 22),$$

e utilizando a mesma chave, temos que o inverso multiplicativo de 5 em $\mathbb{Z}_{26}$ é $n^{-1} = -5$, pois $5 \cdot (-5) = -25 \equiv 1 \pmod{26}$. Logo, a decifragem é dada por:

$$a = (b - 4)(-5) \pmod{26}.$$

Dessa forma, obtemos $(7, 20, 12, 1, 4, 17, 19, 14)$, que corresponde, em letras, à palavra ***HUMBERTO****.*

A *cifra afim* apresenta uma vulnerabilidade importante: ela pode ser decifrada a partir do conhecimento de apenas duas associações entre letras do texto original e do texto cifrado. Em particular, é possível realizar uma *análise de frequência*, associando as letras mais comuns no texto cifrado às letras mais frequentes no idioma — como, no português, as letras "A" e "E". Dessa forma, o problema se reduz à determinação dos parâmetros n e k em $\mathbb{Z}_q$, o que permite reconstruir a chave do sistema.

Definição 105. *Uma cifra de substituição periódica de período p consiste em p alfabetos de cifra $B_1, \ldots, B_p$ e nas chaves $f_i : \mathcal{A} \to B_i$, para $1 \leq i \leq p$.*

Uma mensagem $m = m_1 m_2 \ldots m_p m_{p+1} \ldots m_{2p} \ldots$ é cifrada aplicando-se as funções $f_1, \ldots, f_p$ em sequência, i.e.,

$$m \mapsto f_1(m_1)\, f_2(m_2) \cdots f_p(m_p)\, f_1(m_{p+1})\, f_2(m_{p+2}) \cdots ,$$

repetindo-se o ciclo de funções $f_1, \ldots, f_p$ a cada p caracteres.

A cifra de Vigenère é um dos exemplos mais notórios de cifra de substituição periódica. Nesse método, a chave é composta por uma sequência de letras $k = k_1 k_2 \ldots k_p$.

Cada letra k_i da chave indica o deslocamento cíclico aplicado à i-ésima letra do texto, ou seja, para letra k_i se aplica uma cifra César, com a encriptação sendo dada por:

$$f_i(m) = m + k_i \pmod{26}.$$

Ou seja, cada posição do texto é cifrada com base em um alfabeto diferente, conforme o valor da chave correspondente. Para facilitar o processo de encriptação, utiliza-se frequentemente o quadrado de Vigenère, uma tabela onde cada linha representa uma rotação do alfabeto. Abaixo temos esse quadro de forma simplificada.

	A	**B**	**C**	**...**	**Z**
A	A	B	C	...	Z
B	B	C	D	...	A
C	C	D	E	...	B
...	...	...	...	...	...
Z	Z	A	B	...	Y

Exemplo 106. *Usando a mesma representação numérica das letras adotada no Exemplo 104, consideremos o texto original* ***EXEMPLO****, representado por $(4, 23, 4, 12, 15, 11, 14)$ e a chave* ***SEGREDO****, por $(18, 4, 6, 17, 4, 3, 14)$.*

Na cifra de Vigenère, cada letra do texto é cifrada pela soma módulo 26 de seu valor com o valor correspondente da chave, ou seja,

$$f_i(a_i) = a_i + k_i \pmod{26}.$$

Aplicando a transformação:

$$\begin{aligned} f_1(4) &= 4 + 18 = 22, \\ f_2(23) &= 23 + 4 = 27 \equiv 1, \\ f_3(4) &= 4 + 6 = 10, \\ f_4(12) &= 12 + 17 = 29 \equiv 3, \\ f_5(15) &= 15 + 4 = 19, \\ f_6(11) &= 11 + 3 = 14, \\ f_7(14) &= 14 + 14 = 28 \equiv 2. \end{aligned}$$

Assim, obtemos o texto cifrado

$$b = (22, 1, 10, 3, 19, 14, 2),$$

que corresponde, em letras, à palavra ***WBKDTOC****.*

Outros dois exemplos de *cifras de substituição* sobre $\mathbb{Z}_q$ podem ser obtidos a partir da sequência (f_i), $i = 1, 2, \ldots$, onde cada f_i é aplicada à i-ésima letra a_i do texto original, sendo f_i definida como:

$$f_i : \mathbb{Z}_q \longrightarrow \mathbb{Z}_q, \quad a \mapsto b_i := ka + d_i,$$

com $mdc(k, q) = 1$, e onde

$$d_i = ca_{i-1}, \quad \text{com } c, a_0 \text{ dados},$$

ou alternativamente,

$$d_i = cb_{i-1}, \quad \text{com } c, b_0 \text{ dados}.$$

Exemplo 107. *(i) Utilizando novamente a correspondência numérica entre as letras A a Z apresentada no Exemplo 104, consideremos a palavra* ***EXEMPLO****, isto é:*

$$(4, 23, 4, 12, 15, 11, 14) = a_1a_2a_3a_4a_5a_6a_7,$$

que será cifrada pela função

$$f_i(a) = 3a + a_{i-1},$$

com $a_0 := 4$.

Aplicando a transformação, obtemos:

$$\begin{aligned} f_1(4) &= 3 \cdot 4 + 4 = 16, \\ f_2(23) &= 3 \cdot 23 + 4 = 73 \equiv 21, \\ f_3(4) &= 3 \cdot 4 + 23 = 35 \equiv 9, \\ f_4(12) &= 3 \cdot 12 + 4 = 40 \equiv 14, \\ f_5(15) &= 3 \cdot 15 + 12 = 57 \equiv 5, \\ f_6(11) &= 3 \cdot 11 + 15 = 48 \equiv 22, \\ f_7(14) &= 3 \cdot 14 + 11 = 53 \equiv 1. \end{aligned}$$

Assim, obtemos o texto cifrado:

$$(16, 21, 9, 14, 5, 22, 1),$$

que corresponde, em letras, à palavra ***QVJNFWB****.*

(ii) Mantendo as mesmas condições e a mesma mensagem original ***EXEMPLO****, consideremos agora a cifra definida pelo conjunto de funções:*

$$f_i(a) = 3a + b_{i-1}, \quad \textit{com } b_0 := 4.$$

Dessa forma:

$$\begin{aligned} f_1(4) &= 3 \cdot 4 + 4 = 16, & b_1 &= 16, \\ f_2(23) &= 3 \cdot 23 + 16 = 85 \equiv 7, & b_2 &= 7, \\ f_3(4) &= 3 \cdot 4 + 7 = 19, & b_3 &= 19, \\ f_4(12) &= 3 \cdot 12 + 19 = 55 \equiv 3, & b_4 &= 3, \\ f_5(15) &= 3 \cdot 15 + 3 = 48 \equiv 22, & b_5 &= 22, \\ f_6(11) &= 3 \cdot 11 + 22 = 55 \equiv 3, & b_6 &= 3, \\ f_7(14) &= 3 \cdot 14 + 3 = 45 \equiv 19, & b_7 &= 19. \end{aligned}$$

Assim, o texto cifrado obtido é: $(16, 7, 19, 3, 22, 3, 19)$*, que corresponde, em letras, à palavra* ***HTDWDDT****.*

A cifra de Vigenère foi, durante muitos anos, considerada indecifrável, recebendo inclusive o título de *le chiffre indéchiffrable*. Essa reputação se devia ao fato de que, diferentemente das cifras de substituição simples, cada letra do texto é cifrada segundo um deslocamento que varia periodicamente, o que mascara a frequência natural das letras e dificulta uma análise direta.

No entanto, essa segurança aparente depende fortemente do comprimento da chave. Quando a chave é curta, o processo de encriptação se torna periódico, e o texto cifrado pode ser visto como a justaposição de várias cifras de substituição monoalfabéticas — uma para cada posição da chave. Em termos matemáticos, se o comprimento da chave é p, então cada p-ésima letra do texto é cifrada pela mesma função

$$f_i(a) = a + k_i \pmod{26},$$

com k_i fixo para cada i.

Essa estrutura periódica permite que o criptanalista descubra o valor de p aplicando métodos que exploram repetições de padrões no criptograma. Uma vez conhecido o período, o texto cifrado pode ser dividido em p subsequências independentes, cada uma correspondendo a uma cifra de substituição simples. Assim, pode-se aplicar *análise de*

frequência separadamente a cada subsequência, identificando as letras mais comuns e deduzindo o valor de cada k_i, reconstruindo, portanto, toda a chave.

Dessa forma, a vulnerabilidade da cifra de Vigenère está diretamente ligada à regularidade matemática introduzida pela repetição da chave. Quanto menor o período e mais previsíveis forem os valores k_i, mais fácil se torna quebrar o código por meio de métodos estatísticos. A cifra só se torna realmente segura quando a chave é *tão longa quanto a mensagem* e usada apenas uma vez, princípio que deu origem à *cifra de uso único*, considerada inquebrável.

Definição 108. *Uma* cifra de Hill *com matriz-chave K é uma aplicação de $\left(\mathbb{Z}_q\right)^n$ em si mesma, definida por*

$$a \longmapsto Ka + d,$$

tal que $mdc(\det K, q) = 1$.

Para cifrar, subdivide-se o texto original em blocos de n letras, substitui-se cada letra pelo seu correspondente em $\mathbb{Z}_q$ e escreve-se esse vetor em coluna. Aplica-se então a transformação linear dada por K a cada bloco a.

Exemplo 109. *Consideremos $q = 26$, de modo que os caracteres são representados pelos números $A = 0, B = 1, \ldots, Z = 25$. Utilizemos a matriz-chave*

$$K = \begin{pmatrix} 3 & 2 \\ 5 & 7 \end{pmatrix}, \quad \textit{com } \det(K) = 3 \cdot 7 - 2 \cdot 5 = 11,$$

sendo $mdc(11, 26) = 1$.

Dividimos a palavra ***EXEMPLOS*** *em blocos de duas letras:*

$$(4, 23), \quad (4, 12), \quad (15, 11), \quad (14, 18).$$

Cada bloco $a = (a_1, a_2)^T$ é cifrado pela regra

$$b = Ka \pmod{26}.$$

Calculando:

$$b_1 = K\begin{pmatrix} 4 \\ 23 \end{pmatrix} \equiv \begin{pmatrix} 3 \cdot 4 + 2 \cdot 23 \\ 5 \cdot 4 + 7 \cdot 23 \end{pmatrix} = \begin{pmatrix} 12 + 46 \\ 20 + 161 \end{pmatrix} \equiv \begin{pmatrix} 58 \\ 181 \end{pmatrix} \equiv \begin{pmatrix} 6 \\ 25 \end{pmatrix} \pmod{26},$$

$$b_2 = K\begin{pmatrix} 4 \\ 12 \end{pmatrix} \equiv \begin{pmatrix} 3 \cdot 4 + 2 \cdot 12 \\ 5 \cdot 4 + 7 \cdot 12 \end{pmatrix} = \begin{pmatrix} 36 \\ 104 \end{pmatrix} \equiv \begin{pmatrix} 10 \\ 0 \end{pmatrix} \pmod{26},$$

$$b_3 = K\begin{pmatrix} 15 \\ 11 \end{pmatrix} \equiv \begin{pmatrix} 3 \cdot 15 + 2 \cdot 11 \\ 5 \cdot 15 + 7 \cdot 11 \end{pmatrix} = \begin{pmatrix} 67 \\ 140 \end{pmatrix} \equiv \begin{pmatrix} 15 \\ 10 \end{pmatrix} \pmod{26},$$

$$b_4 = K\begin{pmatrix} 14 \\ 18 \end{pmatrix} \equiv \begin{pmatrix} 3 \cdot 14 + 2 \cdot 18 \\ 5 \cdot 14 + 7 \cdot 18 \end{pmatrix} = \begin{pmatrix} 42 + 36 \\ 70 + 126 \end{pmatrix} = \begin{pmatrix} 78 \\ 196 \end{pmatrix} \equiv \begin{pmatrix} 0 \\ 14 \end{pmatrix} \pmod{26}.$$

Logo, os blocos cifrados são:

$$(6, 25),\ (10, 0),\ (15, 10),\ (0, 14),$$

que correspondem, respectivamente, às letras ***GZKAPKOE****.*

Blocos de duas letras xy sobre um alfabeto com N letras, podem ser representados em correspondência biunívoca pelos N^2 inteiros $xN + y$. A cifra correspondente de dígrafos possui a transformação de encriptação

$$C \equiv aP + b \pmod{N^2},$$

para obter o texto cifrado C a partir do texto original P, onde P e C são dígrafos. De forma análoga, a transformação de decifragem é dada por

$$P \equiv a'C + b' \pmod{N^2}.$$

Exemplo 110. *Seja $n = 26$, $a = 15$ e $b = 49$. O dígrafo* ***PE*** *corresponde ao inteiro*

$$P \cdot n + E = 15 \cdot 26 + 4 = 394.$$

O texto cifrado correspondente é

$$C \equiv aP + b = 15 \cdot 394 + 49 = 551 \pmod{676},$$

portanto, $C = 551$.

A quebra de uma cifra de dígrafos baseia-se na análise da frequência de blocos de duas letras em um texto cifrado (com comprimento razoável), comparando-a com a frequência típica desses blocos em textos comuns de um determinado idioma. Em textos suficientemente longos em português, os dígrafos **DE** e **ES** são os que ocorrem com maior frequência.

Agora apresentaremos as *chaves de permutação*. Seja $\mathbb{F}_q$ um corpo finito com $q = p^n$ elementos, com p primo, e seja ζ um elemento primitivo de $\mathbb{F}_q$. Cada elemento de $\mathbb{F}_q$ pode ser expresso na forma

$$a = c_0\zeta^{n-1} + c_1\zeta^{n-2} + \cdots + c_{n-1}, \quad c_i \in \mathbb{F}_p.$$

O corpo $\mathbb{F}_q$ consiste dos elementos

$$\{\, 0, \zeta, \zeta^2, \ldots, \zeta^{q-1} \,\}.$$

O elemento primitivo ζ é raiz de um polinômio primitivo e irredutível f sobre $\mathbb{F}_p$.

Chamamos o número inteiro

$$c_0p^{n-1} + c_1p^{n-2} + \cdots + c_{n-1}, \quad c_i \in \mathbb{F}_p,$$

de *valor-numérico* do elemento $a \in \mathbb{F}_q$.

Definição 111. *Um polinômio $g \in \mathbb{F}_q[x]$ é chamado de polinômio de permutação de $\mathbb{F}_q$ se a função polinomial $\tilde{g}$ induzida por g é uma permutação de $\mathbb{F}_q$.*

Se $\mathcal{A}$ é um alfabeto com q letras, então uma chave pode ser definida da seguinte forma. Primeiro, estabelece-se uma correspondência biunívoca entre as q letras de $\mathcal{A}$ e os q elementos de $\mathbb{F}_q$. Se $P_1P_2P_3\ldots$ é um texto original, então seja $a_1a_2a_3\ldots$ a sequência correspondente de elementos de $\mathbb{F}_q$. Essa sequência será transformada em

$$\tilde{g}(a_1)\tilde{g}(a_2)\tilde{g}(a_3)\ldots$$

por meio da aplicação do polinômio de permutação g sobre $\mathbb{F}_q$.

Os elementos $\tilde{g}(a_i)$ são então representados como letras de $\mathcal{A}$ e formam o criptograma. Se $f : \mathcal{A} \to \mathbb{F}_q$ é bijetiva e $P \in \mathcal{A}$, então

$$P \mapsto f(P) \mapsto \tilde{g}(f(P)) \mapsto f^{-1}(\tilde{g}(f(P))),$$

é uma chave dada por $f^{-1} \circ \tilde{g} \circ f$.

Se utilizarmos polinômios de permutação g_i no processo de cifragem, obtemos uma chave variável.

Exemplo 112. *Seja $f(x) = x^3 + 2x + 1$ um polinômio primitivo e seja ζ uma raiz de f em $\mathbb{F}_{3^3}$. Os elementos de $\mathbb{F}_{3^3}$ são apresentados na Tabela 112.*

A coluna (1) será utilizada na cifragem de textos. A coluna (2) representa os valores numéricos (field values) dos elementos de $\mathbb{F}_{3^3}$. Na coluna (3), descrevemos esses elementos na forma vetorial (c_0, c_1, c_2), interpretando $\mathbb{F}_{3^3} \cong \mathbb{F}_3[\zeta]/\langle f(\zeta)\rangle$. A coluna (4) apresenta os expoentes das potências ζ^i dos elementos do grupo cíclico $\mathbb{F}_{3^3}^*$. Denotaremos o vetor 000 *pelo expoente* 0*, apesar de $0 \notin \mathbb{F}_{3^3}^*$.*

A chave para a tabela é o fato de que ζ satisfaz

$$\zeta^3 + 2\zeta + 1 = 0,$$

e como toda a aritmética é feita módulo 3*, obtemos a identidade fundamental:*

$$\zeta^3 = \zeta + 2.$$

A correspondência entre colunas (2) e (3) é feita lendo o vetor da coluna (3) como um número em base 3*. Por exemplo, o vetor* 002 *(letra C) representa o elemento $2 \in \mathbb{F}_{3^3}$.*

(1)	*(2)*	*(3)*	*(4)*	*(1)*	*(2)*	*(3)*	*(4)*
A	*0*	*000*	*0*	*O*	*14*	*112*	*11*
B	*1*	*001*	*26*	*P*	*15*	*120*	*4*
C	*2*	*002*	*13*	*Q*	*16*	*121*	*18*
D	*3*	*010*	*1*	*R*	*17*	*122*	*7*
E	*4*	*011*	*9*	*S*	*18*	*200*	*15*
F	*5*	*012*	*3*	*T*	*19*	*201*	*25*
G	*6*	*020*	*14*	*U*	*20*	*202*	*8*
H	*7*	*021*	*16*	*V*	*21*	*210*	*17*
I	*8*	*022*	*22*	*W*	*22*	*211*	*20*
J	*9*	*100*	*2*	*X*	*23*	*212*	*5*
K	*10*	*101*	*21*	*Y*	*24*	*220*	*23*
L	*11*	*102*	*12*	*Z*	*25*	*221*	*24*
M	*12*	*110*	*10*	␣	*26*	*222*	*19*
N	*13*	*111*	*6*				

Assim, para a palavra ***PEDRO*** *temos:*

$$P = 120, \quad E = 011, \quad D = 010, \quad R = 122, \quad O = 112.$$

Seja $g_i = a_i x + b_i$ *com* $a_i \neq 0$*, e defina:*

$$a_{i+2} = a_{i+1}a_i, \qquad b_{i+2} = b_{i+1} + b_i, \qquad i = 1, 2, \ldots,$$

com valores iniciais do exemplo:

$$a_1 = 021 = \zeta^{16}, \qquad a_2 = 111 = \zeta^{6}, \qquad b_1 = 002, \qquad b_2 = 110.$$

A cifragem da palavra ***PEDRO*** *é mostrada a seguir:*

(a)	P	E	D	R	O
(b)	120	011	010	122	112
(c)	4	9	1	7	11
(d)	16	6	22	2	24
(e)	211	200	220	011	011
(f)	002	110	112	222	001
(g)	210	010	002	200	012
(h)	V	D	C	S	F

A leitura das linhas é a seguinte:

(a) texto original;

(b) forma vetorial dos símbolos em $\mathbb{F}_{3^3}$*;*

(c) expoentes t *correspondentes a* $\alpha_i = \zeta^t$ *conforme a tabela;*

(d) expoentes de a_i*;*

(e) primeiro somamos $(c) + (d)$ *módulo* 26*; em seguida consultamos a tabela do Exemplo 112 para recuperar o elemento correspondente, escrevendo-o finalmente na forma vetorial;*

(f) forma vetorial de b_i*;*

(g) forma vetorial de $y_i = a_i\alpha_i + b_i = (e) + (f)$*;*

(h) texto cifrado obtido.

4.1.2 Criptografia de chave pública e RSA

A criptografia de chave pública surgiu a partir da observação de que o processo e a chave usados para cifrar uma mensagem não precisam ser os mesmos usados para decifrá-la. Isso elimina a necessidade de que todas as chaves sejam mantidas em segredo ou distribuídas por canais seguros, pois a chave de cifragem pode ser tornada pública, enquanto apenas a chave privada permanece protegida.

Esse modelo facilita a gestão de chaves em sistemas com muitos usuários e ainda permite a implementação de assinaturas digitais e mecanismos de autenticação. Em um cenário típico, mesmo que um hacker tenha acesso ao texto cifrado e à chave pública, ele não consegue recuperar a chave privada necessária para decifrar a mensagem.

Em sistemas de criptografia de chave pública, os métodos de cifragem E e decifragem D devem satisfazer:

1. **Correção:** para toda mensagem M,

$$D(E(M)) = M,$$

   e sem conhecer D deve ser praticamente impossível reverter E.

2. **Eficiência:** os algoritmos E e D devem ser fáceis de computar.

3. **Assimetria:** o método E pode ser público sem permitir deduzir D.

4. **Reversibilidade:** também deve valer

$$E(D(M)) = M,$$

   propriedade utilizada em mecanismos como assinaturas digitais.

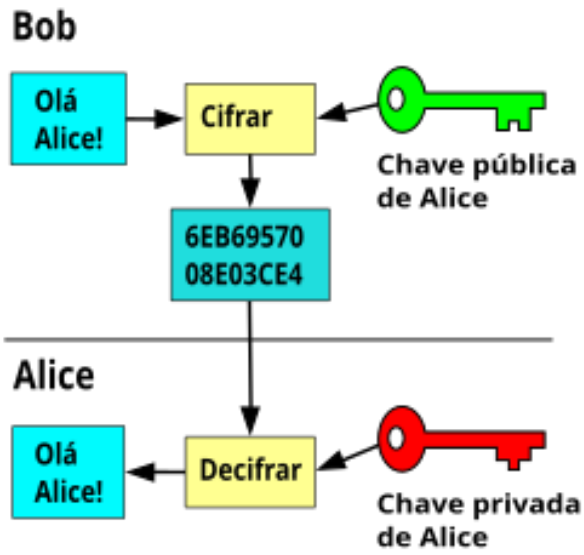

Figura 2 – Diagrama de Criptografia de chave pública. Acessado em https://pt.wikipedia.org/wiki/Criptografia_de_chave_p%C3%BAblica

No uso prático desses sistemas de chave pública, os participantes seguem um protocolo. Seja B um usuário que gera suas chaves E_B (pública) e D_B (privada). A chave E_B é disponibilizada em um diretório público. Quando uma usuária A deseja enviar uma mensagem privada a B, ela aplica E_B à mensagem; somente B, que possui D_B, é capaz de decifrá-la. A segurança do processo depende da autenticidade de E_B; se uma chave falsa for publicada por um impostor, o sigilo da comunicação é perdido. Por isso, na prática utiliza-se frequentemente uma autoridade de gerenciamento de chaves.

Uma vantagem adicional dos criptossistemas de chave pública é a possibilidade de autenticação por meio de *assinaturas digitais*. Para assinar uma mensagem M, o usuário B calcula

$$S = D_B(M),$$

utilizando sua chave privada. Em seguida, envia $E_A(S)$ ao destinatário A, que pode recuperar a assinatura aplicando seu método de decifragem D_A. Depois, A utiliza a chave pública de B para verificar que

$$E_B(S) = M.$$

Como somente B possui sua chave privada, A tem garantia de que a mensagem é autêntica. A probabilidade de um valor aleatório produzir uma mensagem significativa após a verificação é extremamente pequena, o que impede falsificações.

A construção desses sistemas depende de *funções de mão única* (one-way functions): funções fáceis de calcular, mas cujo inverso é computacionalmente impraticável de determinar. Quando a inversão se torna fácil mediante informação secreta (a chave de decifragem), diz-se que a função é de mão única com *atalho* (trapdoor). A ausência de estrutura simples nessas funções é essencial para a segurança dos criptossistemas.

Relembremos que a função φ de Euler indica a quantidade de inteiros no conjunto $\{0, 1, \ldots, n-1\}$ que são coprimos com n. Em particular, para um número primo p, tem-se $\varphi(p) = p - 1$. Além disso, se p e q são primos distintos, então

$$\varphi(pq) = (p-1)(q-1).$$

Um resultado central da teoria dos números é o *Teorema de Euler*, o qual afirma que, para qualquer inteiro a relativamente primo a n,

$$a^{\varphi(n)} \equiv 1 \pmod{n}.$$

Esse teorema generaliza o Pequeno Teorema de Fermat e desempenha papel essencial na construção de diversos criptossistemas modernos.

A demonstração sai como consequencia dos Teoremas 10 e 11.

Com base nesses resultados Rivest, Shamir e Adleman (1978) introduziram o criptossistema RSA, cuja segurança permanece confiável.

Definição 113 (Criptossistema RSA)**.** *1. Selecione dois números primos grandes p e q, e defina*

$$n = pq.$$

2. Escolha um inteiro aleatório grande d tal que $mdc(d, \varphi(n)) = 1$. Observe que

$$\varphi(n) = (p-1)(q-1).$$

3. Determine $e \in \mathbb{Z}$, com $1 \leq e < \varphi(n)$, tal que

$$ed \equiv 1 \pmod{\varphi(n)}.$$

Além disso, mdc(e, $\varphi(n)$) = 1.

4. Torne tanto n quanto e públicos (chave pública) e mantenha d e os primos p e q em segredo (chave privada).

5. Represente o texto original como uma sequência de inteiros m, com $0 \leq m \leq n-1$.

6. Cifragem:

$$c \equiv m^e \pmod{n}.$$

7. Decifragem:

$$m \equiv c^d \pmod{n}.$$

Teorema 114 (Teorema Chinês do Resto)**.** *Sejam $m_1, \ldots, m_r$ inteiros positivos tais que $mdc(m_i, m_j) = 1$ para todo $i \neq j$. Dados quaisquer inteiros $a_1, \ldots, a_r$, o sistema de congruências*

$$\begin{cases} x \equiv a_1 \pmod{m_1}, \\ x \equiv a_2 \pmod{m_2}, \\ \quad \vdots \\ x \equiv a_r \pmod{m_r}, \end{cases}$$

possui solução $x \in \mathbb{Z}$. Além disso, todas as soluções são congruentes entre si módulo $M := m_1 m_2 \cdots m_r$ (isto é, a solução é única módulo M).

Demonstração. Mostraremos primeiro a existência de solução e, em seguida, sua unicidade módulo $m_1 m_2 \cdots m_r$. A demonstração será feita por indução sobre r.

Para $r = 2$.

Considere o sistema

$$\begin{cases} x \equiv a \pmod{m}, \\ x \equiv b \pmod{n}, \end{cases}$$

onde $m, n \in \mathbb{Z}_{>0}$ e $mdc(m, n) = 1$. Mostraremos que existe $x \in \mathbb{Z}$ que satisfaça ambas as congruências, e depois provaremos que tal solução é única módulo mn.

Escrevemos a primeira congruência como uma equação em $\mathbb{Z}$:

$$x = a + my \qquad \text{para algum } y \in \mathbb{Z}.$$

Substituímos esta expressão na segunda congruência:

$$a + my \equiv b \pmod{n}.$$

Subtraindo a dos dois lados, obtemos

$$my \equiv b - a \pmod{n}. \tag{1}$$

Como $mdc(m, n) = 1$, o inteiro m é invertível módulo n; seja m_0 um inverso de m módulo n, isto é,

$$mm_0 \equiv 1 \pmod{n}.$$

Multiplicando a congruência (1) por m_0, encontramos:

$$y \equiv m_0(b - a) \pmod{n}.$$

Logo,

$$y = m_0(b - a) + nz, \qquad z \in \mathbb{Z}.$$

Substituindo em $x = a + my$, obtemos a forma geral das soluções:

$$x = a + m(m_0(b-a) + nz) = a + mm_0(b-a) + mnz.$$

Verifiquemos que esse x satisfaz ambas as congruências.

Primeiro:

$$x \equiv a + 0 + 0 \equiv a \pmod{m}.$$

Segundo:

$$x \equiv a + 1(b-a) + 0 \equiv b \pmod{n}.$$

Assim, existe solução para $r = 2$.

Suponha o resultado válido para qualquer sistema com r módulos dois a dois coprimos. Considere então $r+1$ módulos $m_1, \ldots, m_r, m_{r+1}$ dois a dois coprimos e os inteiros $a_1, \ldots, a_{r+1}$. Pelo passo indutivo, existe $b \in \mathbb{Z}$ que resolve o sistema das primeiras r congruências:

$$b \equiv a_i \pmod{m_i} \quad \text{para } i = 1, \ldots, r.$$

Considere agora o sistema de duas congruências

$$\begin{cases} x \equiv b \pmod{m_1 m_2 \cdots m_r}, \\ x \equiv a_{r+1} \pmod{m_{r+1}}. \end{cases}$$

Como cada m_i (com $1 \leq i \leq r$) é coprimo a m_{r+1}, segue que $mdc(m_1 \cdots m_r, m_{r+1}) = 1$. Pelo caso $r = 2$, existe c que simultaneamente satisfaça as duas congruências acima. Então, para todo $i = 1, \ldots, r$, temos $c \equiv b \equiv a_i \pmod{m_i}$ e, por construção, $c \equiv a_{r+1} \pmod{m_{r+1}}$. Portanto c resolve todas as $r+1$ congruências, provando a existência para $r+1$. Pelo princípio da indução, o sistema tem solução para todo r

Agora provaremos a unicidade. Sejam x e x' duas soluções do sistema com módulos $m_1, \ldots, m_r$. Então $x \equiv x' \pmod{m_i}$ para cada i, ou seja $m_i \mid (x - x')$. Como os m_i são mutuamente primos, o produto $M = m_1 \cdots m_r$ divide $x - x'$. Assim $x \equiv x' \pmod{M}$. Logo todas as soluções são idênticas módulo M, o que garante a unicidade da solução até essa equivalência. □

Teorema 115. *Dadas a chave pública $\{e, n\}$ e a chave privada d (e os primos p e q) de um criptossistema RSA, então, para qualquer mensagem m com $0 \leq m \leq n-1$, a transformação de cifragem é*

$$E(m) = c \equiv m^e \pmod{n},$$

e a transformação de decifragem é

$$D(c) = c^d \equiv m \pmod{n},$$

isto é, $D(E(m)) \equiv m \pmod{n}$.

Demonstração. Seja $c \equiv m^e \pmod{n}$. Então,

$$D(c) \equiv c^d \equiv (m^e)^d \equiv m^{ed} \pmod{n}.$$

Como $ed \equiv 1 \pmod{\varphi(n)}$, existe um inteiro k tal que

$$ed = 1 + k\varphi(n).$$

Assim,

$$m^{ed} = m^{1+k\varphi(n)} = m \cdot m^{k\varphi(n)} = m \cdot (m^{\varphi(n)})^k.$$

Caso 1: $mdc(m, n) = 1$

Pelo Teorema de Euler, se $mdc(m, n) = 1$,

$$m^{\varphi(n)} \equiv 1 \pmod{n}.$$

Portanto,

$$m^{ed} \equiv m \cdot (m^{\varphi(n)})^k \equiv m \cdot (1)^k \equiv m \pmod{n}.$$

Caso 2: $mdc(m, n) \neq 1$

Se $mdc(m, n) \neq 1$, como $n = p \cdot q$ (p e q primos), m deve ser um múltiplo de p ou um múltiplo de q (ou ambos). Precisamos mostrar que $m^{ed} \equiv m \pmod{p}$ e $m^{ed} \equiv m \pmod{q}$.

Como $ed \equiv 1 \pmod{\varphi(n)}$ e $(p-1)$ divide $\varphi(n)$, temos $ed \equiv 1 \pmod{p-1}$. Seja $ed = 1 + j(p-1)$ para algum inteiro j.

- Se $m \equiv 0 \pmod{p}$ (p divide m), então $m^{ed} \equiv 0^{ed} \equiv 0 \pmod{p}$. Portanto, $m^{ed} \equiv m \pmod{p}$.
- Se $m \not\equiv 0 \pmod{p}$ ($mdc(m, p) = 1$), pelo Pequeno Teorema de Fermat ($m^{p-1} \equiv 1 \pmod{p}$):

$$m^{ed} = m^{1+j(p-1)} = m \cdot (m^{p-1})^j \equiv m \cdot 1^j \equiv m \pmod{p}.$$

Em ambos os subcasos, $m^{ed} \equiv m \pmod{p}$.

De forma análoga, como $ed \equiv 1 \pmod{q-1}$, podemos provar em todos os subcasos que:

$$m^{ed} \equiv m \pmod{q}.$$

Pelo Teorema Chinês do Resto (TCR), já que $m^{ed} \equiv m$ é válido tanto módulo p quanto módulo q, e $mdc(p, q) = 1$, a congruência é válida módulo o produto $n = pq$:

$$m^{ed} \equiv m \pmod{n}.$$

□

Um aspecto fundamental do criptossistema RSA é que o valor n utilizado como módulo é o produto de dois números primos grandes, aproximadamente do mesmo tamanho. A segurança do método depende do fato de que fatorar inteiros muito grandes é computacionalmente inviável com a tecnologia atual.

Durante o processo de cifragem, a mensagem é dividida em blocos, e cada bloco é interpretado como um número no intervalo

$$0 \leq m \leq n-1,$$

uma vez que todas as operações do criptossistema são realizadas módulo n. O ponto crucial é que, quando n é suficientemente grande, torna-se praticamente impossível determinar o expoente de decifragem d, que satisfaz

$$ed \equiv 1 \pmod{\varphi(n)},$$

sem conhecer previamente a fatoração de n nos primos p e q. Somente o receptor autorizado possui esses valores e, portanto, é capaz de calcular

$$\varphi(n) = (p-1)(q-1)$$

e determinar o inverso multiplicativo de e módulo $\varphi(n)$.

Para determinar o expoente de decifragem d no RSA, é necessário que d seja o inverso multiplicativo de e módulo $\varphi(n)$, isto é,

$$ed \equiv 1 \pmod{\varphi(n)} \Rightarrow ed = 1 + k\varphi(n) \Rightarrow ed + y\varphi(n) = 1,\ k \in \mathbb{Z}$$

A obtenção desse inverso é feita por meio do algoritmo de Euclides estendido, que resolve equações diofantinas da forma

$$ax + by = mdc(a, b).$$

O algoritmo consiste em:

Executam-se as divisões sucessivas(algoritmo usual)

$$\begin{aligned} \varphi(n) &= q_1 e + r_1, \\ e &= q_2 r_1 + r_2, \\ r_1 &= q_3 r_2 + r_3, \\ &\vdots \end{aligned}$$

até que o último resto seja 1, confirmando que $mdc(e, \varphi(n)) = 1$.

Agora escrevemos o último resto como combinação linear dos restos anteriores:

$$1 = r_k = r_{k-2} - q_k r_{k-1},$$

e substituímos recursivamente cada r_i pelas expressões obtidas nas divisões anteriores, até obter

$$1 = ed + \varphi(n)y.$$

Caso d seja negativo, escolhe-se o representante positivo adicionando-se $\varphi(n)$.

Esse processo é eficiente apenas quando se conhece $\varphi(n)$, isto é, quando os primos p e q são conhecidos. Para um hacker descobrir $\varphi(n)$ é necessário fatorar n, operação considerada computacionalmente inviável para valores suficientemente grandes de n.

Nos exemplos a seguir consideraremos $A = 01, B = 02, \ldots, Z = 26$ e "espaço"$= 00$.

Exemplo 116. *Seja $n = pq = 3 \cdot 11 = 33$ e escolha $d = 7$. Então $\varphi(n) = 20$, e a congruência $7e \equiv 1 \pmod{20}$ tem como solução $e = 3$.*

*Para cifrar a palavra **HUMBERTO**, representamo-la em forma numérica, isto é:*

$$08, 21, 13, 02, 05, 18, 20, 15$$

Calculando $8^3 \equiv 17 \pmod{33}$, $21^3 \equiv 21 \pmod{33}$, etc., obtemos o criptograma

$$17, 21, 19, 8, 26, 24, 14, 09$$

*que corresponde à palavra **QUSHZXNI**.*

*Para decifrar o criptograma **LZNXA**, ou*

$$12, 26, 14, 24, 01$$

*calculamos $12^7 \equiv 12 \pmod{33}$, etc., e assim recuperamos a mensagem **LETRA**.*

Exemplo 117. *Seja $n = pq = 47 \cdot 59 = 2773$. Escolha $d = 157$ e utilize o Algoritmo de Euclides Estendido para resolver*

$$157e \equiv 1 \pmod{2668},$$

onde $\varphi(n) = (p-1)(q-1) = 2668$, obtendo-se $e = 17$.

*A mensagem **ONIBUS EM MARROCOS** deve ser dividida em blocos de pelo menos duas letras (ou letras individuais). Assim, os blocos da mensagem são:*

$$1514,\ 0902,\ 2119,\ 0005,\ 1300,\ 1301,\ 1818,\ 1503,\ 1519$$

onde $0 \leq m_i \leq n-1$. Calculando

$$c_i \equiv m_i^e \pmod{n},$$

obtemos os criptogramas:

$$2214,\ 1602,\ 0967,\ 0508,\ 0446,\ 2556,\ 1992,\ 1832,\ 1204$$

Para decifrar os números cifrados $2214, 1602, \ldots$*, temos por exemplo*

$$2214^{157} \equiv 1514 \pmod{2773},\ 1602^{157} \equiv 902 \pmod{2773},\ \cdots$$

Como os blocos possuem quatro dígitos, concluímos que 1514 *corresponde à mensagem* ***ON****.*

Destacam-se alguns aspectos fundamentais relacionados à segurança do sistema RSA. Em primeiro lugar, embora o cálculo de $\varphi(n)$ não seja, em geral, mais simples do que a fatoração de n, o conhecimento de $\varphi(n)$ torna possível recuperar diretamente os fatores primos de n. De fato,

$$\varphi(n) = (p-1)(q-1) = n - (p+q) + 1,$$

de modo que, se n e $\varphi(n)$ forem conhecidos, obtém-se a soma $p+q$. Além disso,

$$(p-q)^2 = (p+q)^2 - 4n,$$

o que permite determinar também $p-q$. Com essas duas relações, os fatores podem ser recuperados por

$$p = \tfrac{1}{2}\big[(p+q) + (p-q)\big], \qquad q = \tfrac{1}{2}\big[(p+q) - (p-q)\big].$$

Portanto, conhecer $\varphi(n)$ é tão comprometedor quanto conhecer a fatoração de n.

Outra consideração essencial é a escolha de primos seguros para p e q. Esses são primos tais que $p-1$ e $q-1$ possuam grandes fatores primos, e que esses fatores por sua vez também possuam fatores primos grandes. Tal escolha evita ataques por técnicas de iteração, nos quais se analisa o ciclo gerado pela aplicação repetida da transformação de cifragem $c \mapsto c^e \pmod{n}$. Primos seguros garantem ciclos longos, dificultando esse tipo de ataque.

Os primos p e q devem ainda possuir tamanhos semelhantes, mas com números de dígitos distintos. Caso contrário, o método de fatoração de Fermat pode ser utilizado de forma eficiente, pois

$$n = a^2 - b^2 = (a+b)(a-b),$$

e, quando p e q são muito próximos, n está próximo de um quadrado perfeito, reduzindo drasticamente o esforço de busca.

Outra vulnerabilidade surge quando o mesmo módulo n é utilizado por mais de dois usuários. Se dois usuários possuem chaves públicas (e_1, n) e (e_2, n) e recebem a mesma mensagem m, então

$$c_1 \equiv m^{e_1} \pmod{n}, \qquad c_2 \equiv m^{e_2} \pmod{n}.$$

Se $mdc(e_1, e_2) = 1$, existem inteiros a e b tais que

$$ae_1 + be_2 = 1.$$

Nesse caso, o hacker pode recuperar a mensagem calculando

$$m \equiv (c_1^{-1})^{|a|} c_2^b \pmod{n}.$$

Como a probabilidade de que dois expoentes escolhidos aleatoriamente sejam coprimos é alta (aproximadamente 0.8106 no RSA), o uso de um módulo compartilhado é considerado inseguro.

Essas observações evidenciam que a segurança do RSA depende não apenas do tamanho do módulo n, mas também da escolha criteriosa dos primos p e q, dos expoentes e das práticas de implementação utilizadas.

A função de mão única com atalho (*trapdoor one-way*) do RSA é simplesmente a exponenciação discreta

$$c = m^e \pmod{n},$$

no qual o "atalho"(*trapdoor*) é o conjunto $\{p, q, e\}$. A função inversa é dada por

$$m = c^d \pmod{n},$$

onde $de \equiv 1 \pmod{\varphi(n)}$. Tanto o domínio quanto a imagem dessa função são o conjunto dos inteiros entre 0 e $n-1$. Assim, o RSA também pode ser usado para formar assinaturas digitais, conforme descrito anteriormente, logo após a Figura 2. Essa capacidade de produzir assinaturas digitais é uma das características mais úteis do sistema RSA.

Sejam n_A, d_A, e_A os parâmetros RSA da usuária A, onde D_A e E_A denotam, respectivamente, seus algoritmos de decifragem e cifragem; e definimos de forma análoga n_B, d_B, e_B para o usuário B. Suponha que A deseje enviar uma mensagem assinada m para B. O protocolo correspondente é o seguinte:

1. A calcula a assinatura

   $$s = D_A(m) = m^{d_A} \pmod{n_A},$$

   e então cifra essa assinatura utilizando a chave pública de B, cujo par é $\{e_B, n_B\}$. Se $n_B > n_A$, A calcula

   $$c = E_B(s) = s^{e_B} \pmod{n_B},$$

   e envia c para B.

2. B recebe c e aplica seu algoritmo de decifragem:

   $$D_B(c) = D_B(s^{e_B}) = s^{e_B d_B} = s \pmod{n_B}.$$

   Em seguida, usando a chave pública de A, forma

   $$E_A(s) = E_A(D_A(m)) = m^{e_A d_A} = m \pmod{n_A}.$$

Com isso B tem a garantia de que a mensagem m veio de A, pois apenas A poderia ter gerado corretamente s, e B consegue recuperá-la usando sua chave secreta d_B e a chave pública de A.

Se $n_B < n_A$, então A divide s em blocos menores do que n_B e cifra cada bloco separadamente usando a chave pública de B. Alternativamente, se $n_A > n_B$, a mensagem pode ser assinada formando-se $D_A(E_B(m))$.

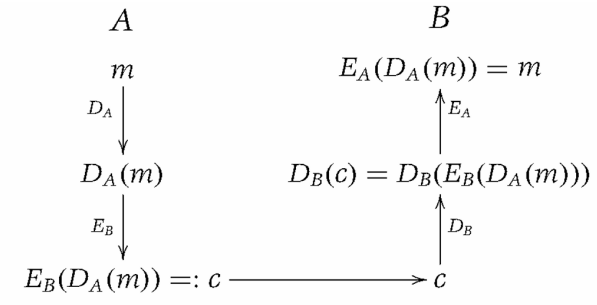

Figura 3 – Diagrama de Criptografia de chave pública. Presente no livro (LIDL; PILZ, 1997)

Exemplo 118. *Suponha que A deseja enviar para B a mensagem em texto puro $m = 207$, assinada pela assinatura digital RSA. As respectivas chaves RSA são dadas por*

$$A\text{: } n_A = 23 \cdot 47 = 1081, \qquad d_A = 675,$$

$$B\text{: } n_B = 31 \cdot 59 = 1829, \qquad e_B = 7.$$

Observamos que $n_B > n_A$. A usuária A calcula

$$s \equiv m^{d_A} \pmod{n_A},$$

isto é,

$$276 \equiv 207^{675} \pmod{1081},$$

e então

$$276^7 \equiv 386 \pmod{1829}.$$

Portanto, a mensagem assinada é 386.

O usuário B recupera a mensagem aplicando primeiro D_B e depois E_A.

4.1.3 Logaritmo discreto e algumas cifras

O logaritmo discreto desempenha um papel fundamental em diversos criptossistemas modernos. Em um corpo finito $\mathbb{F}_q$, fixado um gerador α do grupo multiplicativo $\mathbb{F}_q^*$, o *logaritmo discreto* de um elemento $a \in \mathbb{F}_q^*$ é o único inteiro r, com $0 \leq r < q-1$, tal que

$$a = \alpha^r.$$

Denotamos esse valor por $\log_\alpha(a)$ ou simplesmente $\log(a)$ quando não houver ambiguidade quanto à base. O logaritmo discreto compartilha propriedades semelhantes às do logaritmo real, adaptadas ao contexto modular:

$$\log_\alpha(ab) \equiv \log_\alpha(a) + \log_\alpha(b) \pmod{q-1},$$

$$\log_\alpha(ab^{-1}) \equiv \log_\alpha(a) - \log_\alpha(b) \pmod{q-1}.$$

A operação inversa, a *exponenciação discreta*, é dada por

$$\exp_\alpha(r) = \alpha^r.$$

A exponenciação discreta é facilmente computada por algoritmos eficientes, que permite calcular α^r utilizando apenas um número reduzido de multiplicações. Por outro lado, calcular o logaritmo discreto — isto é, determinar r a partir de α^r — é considerado computacionalmente difícil para corpos apropriados, constituindo assim uma função de mão única. Essa assimetria é essencial para a segurança de diversos criptossistemas.

A segurança de diversos esquemas criptográficos — tanto clássicos quanto de chave pública — repousa exatamente sobre a dificuldade do problema do logaritmo discreto. Em sistemas baseados em exponenciação discreta, como $c = m^e$, descobrir a chave secreta equivale a resolver a congruência

$$e\log(m) \equiv \log(c) \pmod{q-1},$$

o que, na prática, reduz-se ao problema de calcular logaritmos discretos. Por esse motivo, escolhem-se corpos $\mathbb{F}_q$ tais que $q-1$ possua um grande fator primo, dificultando a aplicação de algoritmos eficientes de ataque.

Atualmente, acredita-se que, para corpos primos $\mathbb{F}_p$, o problema do logaritmo discreto possui dificuldade comparável à da fatoração de inteiros de tamanho semelhante, o que reforça sua relevância como base para construção de protocolos seguros.

Exemplo 119. *Descrevemos um criptossistema clássico, também chamado* algoritmo sem chave *ou* algoritmo de três passos*, para transmissão de mensagens. O usuário A deseja enviar uma mensagem $m \in \mathbb{F}_q^*$ para o usuário B. O corpo finito $\mathbb{F}_q$ é publicamente conhecido.*

A escolhe um inteiro aleatório a, onde $1 \leq a \leq q-1$ e $mdc(a, q-1) = 1$, e envia a mensagem cifrada

$$x = m^a \pmod{q}$$

para B.

Em seguida, B escolhe um inteiro aleatório b, onde $1 \leq b \leq q-1$ e $mdc(b, q-1) = 1$, e envia

$$y = x^b = m^{ab} \pmod{q}$$

para A.

Agora A calcula

$$z = y^{a'} = m^{aba'} \quad (mod\ q),$$

onde a' é o inverso multiplicativo de a módulo $q-1$, isto é,

$$aa' \equiv 1 \pmod{q-1},$$

e envia z para B.

Então B computa

$$m = z^{b'} \quad (mod\ q),$$

onde b' é o inverso multiplicativo de b módulo $q-1$. De fato,

$$((m^{ab})^{a'})^{b'} = m.$$

Como exemplo numérico desta cifra com o corpo finito , $\mathbb{F}_{2^3}$, consideramos a transmissão da mensagem

$$m = 011 = x + x^2.$$

Aqui $\mathbb{F}_{2^3}$ é o anel de polinômios quocientado pelo ideal gerado pelo polinômio irredutível.

$$f(x) = x^3 + x + 1.$$

Suponha que A escolha $a = 13$, compute

$$m^a = (x + x^2)^{13} = 1 + x \quad em\ \mathbb{F}_{2^3},$$

e envie $1 + x$ para B.

O usuário B escolhe $b = 11$, então

$$(1 + x)^{11} = 1 + x + x^2,$$

e envia isso para A.

A encontra $a' = 6$ a partir de $13a' \equiv 1 \pmod{7}$ e envia x^2, que obteve de

$$(1 + x + x^2)^6 = x^2,$$

para B.

Finalmente, B calcula $b' = 2$ a partir de $11b' \equiv 1 \pmod{7}$ e obtém

$$(x^2)^2 = x + x^2 = m.$$

Exemplo 120. *Descrevemos a seguir o* criptossistema de chave pública de El Gamal, *utilizado para a transmissão de mensagens no corpo finito* $\mathbb{F}_q$. *Considere um elemento primitivo* $\alpha \in \mathbb{F}_q$, *sendo tanto* α *quanto* q *de conhecimento público.*

O usuário B inicia o sistema escolhendo um inteiro secreto b *de forma aleatória e publica o valor* $\alpha^b \in \mathbb{F}_q$, *que constitui sua chave pública. O valor* b *permanece em sigilo.*

Quando o usuário A deseja enviar uma mensagem $m \in \mathbb{F}_q^*$, *ele seleciona aleatoriamente um inteiro* a *com* $1 \leq a \leq q-1$ *e envia a B o par*

$$\left(\alpha^a,\, m\,\alpha^{ab}\right).$$

Como B conhece seu valor secreto b, *ele calcula*

$$\alpha^{ab} = (\alpha^a)^b,$$

e, dividindo o segundo componente do par por esse valor, recupera a mensagem original m.

Para um exemplo numérico, considere $q = 71$ *e o elemento primitivo* $\alpha = 7$. *Suponha que o usuário B gere sua chave pública como* $\alpha^b = 3$, *para algum inteiro secreto* b. *O usuário A escolhe o valor aleatório* $a = 2$ *e deseja transmitir a mensagem* $m = 20$.

O par enviado por A é então

$$(\alpha^a,\, m\,\alpha^{ab}) = (49,\, 38).$$

De fato, como $\alpha^{ab} = 9$ *em* $\mathbb{F}_{71}$, *temos:*

$$m\,\alpha^{ab} = 20 \cdot 9 = 38 \pmod{71}.$$

Assim, o usuário B, ao computar $(\alpha^a)^b = \alpha^{ab} = 9$, *pode isolar e recuperar* m *a partir do segundo componente do par recebido.*

Se, no criptossistema El Gamal, a mensagem m é uma chave escolhida por A, então este sistema serve como um mecanismo de troca de chaves. No entanto, não é uma troca de chaves autenticada, pois um impostor poderia se passar por A e enviar mensagens falsas para B.

Exemplo 121. *Seja* p *um número primo e seja* α *um elemento primitivo de* $\mathbb{F}_p$. *Tanto* p *quanto* α *são públicos.*

Suponha que o usuário A deseje enviar uma mensagem assinada para o usuário B. Para isso, A escolhe um inteiro aleatório a, *com* $1 \leq a \leq p-2$, *e calcula*

$$y = \alpha^a \in \mathbb{F}_p.$$

Dado um valor de mensagem m, onde $0 \leq m \leq p-1$, o usuário A escolhe um inteiro secreto k, com $0 \leq k \leq p-1$ e $mdc(k, p-1) = 1$. Em seguida, calcula

$$r = \alpha^k \quad e \quad s = k^{-1}(m - ra) \pmod{p-1}.$$

A então envia para B o triplo composto pela mensagem m, pelo valor r e pela assinatura s.

Para verificar a autenticidade, o usuário B calcula

$$c = \alpha^m \quad e \quad c' = y^r \cdot r^s \in \mathbb{F}_p.$$

Observamos que

$$m \equiv ks + ra \pmod{p-1},$$

de modo que, se a assinatura é válida, então

$$c = \alpha^m = (\alpha^a)^r \cdot (\alpha^k)^s = y^r \cdot r^s = c'.$$

Notamos também que o valor k deve ser diferente para cada mensagem m, pois reutilizar o mesmo k permite que o segredo a seja determinado de forma única. Além disso, é necessário que $p-1$ possua um fator primo grande; caso contrário, o problema do logaritmo discreto — necessário para obter a a partir de α^a, pode ser resolvido de forma eficiente, comprometendo a segurança do sistema.

A dificuldade de encontrar logaritmos discretos no corpo finito $\mathbb{F}_q$ é o que fundamenta muitas das cifras modernas. No caso específico em que $q-1$ possui apenas fatores primos pequenos, o algoritmo **Silver-Pohlig-Hellman** permite calcular de forma eficiente estes logaritmos. Resolver o problema do logaritmo discreto significa: dado um elemento $a \in \mathbb{F}_q^*$ e um elemento primitivo α de $\mathbb{F}_q$, encontrar o valor r para o qual $a = \alpha^r$. Este valor é denotado como $r = \log_\alpha a$. A estratégia central do algoritmo baseia-se na fatoração de $q-1$. Seja a fatoração em primos distintos $q-1 = \prod_{i=1}^{k} p_i^{e_i}$, onde $p_1 < p_2 < \ldots < p_k$. O valor de r será determinado módulo $p_i^{e_i}$ para cada $i = 1, 2, \ldots, k$, e os resultados são combinados usando o Teorema Chinês dos Restos para inteiros, a fim de obter o valor final de $r \pmod{q-1}$, que é único no intervalo $0 \leq r < q-1$.

Para aplicar o teorema, primeiro consideramos a representação de r em termos de p_i. O objetivo é determinar os coeficientes r_j, para $j = 0, \ldots, e_i - 1$, para cada fator primo p_i, com $0 \leq r_j \leq p_i - 1$. A representação é dada por:

$$r \equiv \sum_{j=0}^{e_i-1} r_j p_i^j \pmod{p_i^{e_i}}$$

Os resultados são então combinados resolvendo-se um sistema de congruências para $r \pmod{p_i^{e_i}}$. O primeiro passo é determinar r_0, o que é feito formando-se a seguinte relação:

$$a^{(q-1)/p_i} \equiv \alpha^{(q-1)r/p_i} = c_i^r = c_i^{r_0}$$

Nesta equação, $c_i = \alpha^{(q-1)/p_i}$ é definido como uma raiz p_i-ésima primitiva da unidade. Como existem apenas p_i valores possíveis para $a^{(q-1)/p_i}$ (correspondentes a $r_0 = 0, 1, \ldots, p_i - 1$), a avaliação é computacionalmente viável se p_i for pequeno. Construindo uma tabela com os valores $1, c_i, \ldots, c_i^{p_i-1}$. Ao calcular $a^{(q-1)/p_i}$ e buscar o resultado nesta tabela, o valor de r_0 é determinado unicamente.

Após encontrar r_0, o próximo dígito, r_1, é obtido a partir da expressão $a\alpha^{-r_0} = \alpha^{t_1}$, onde $t_i = \sum_{j=1}^{e_i-1} r_j p_i^j$. Em seguida, a relação

$$\alpha^{t_1(q-1)/p_i^2} = c_i^{t_1/p_i} = c_i^{r_1}$$

determina unicamente o valor de r_1. O processo é repetido para os demais dígitos r_j, etc.

Pode ser demonstrado que este algoritmo possui um tempo de execução da ordem de, no máximo, $p_k^{1/2}(\log q)^2$, onde p_k é o maior fator primo de $q - 1$.

Exemplo 122. *Seja $q = 181$ e $\alpha = 2$. Então $q - 1 = 180 = 2^2 \cdot 3^2 \cdot 5$. Calculamos o logaritmo discreto r de $a = 62$ na base 2. Como $p_1 = 2$, $p_2 = 3$ e $p_3 = 5$, consideramos três partes do algoritmo de Silver-Pohlig-Hellman e combinamos os resultados usando o Teorema Chinês do Resto para inteiros.*

Primeiro determinamos $r = r_0 + 2r_1 \pmod 4$. De $\alpha^{(q-1)/2} = 2^{90} \equiv 180$ e $a^{(q-1)/2} = 62^{90} \equiv 1$ concluímos que $r_0 = 0$. Em seguida, $62 \cdot 2^{-r_0} = 62$ e $62^{45} = 2^{r_1(q-1)/2} = 1$ se $r_1 = 0$, logo $r_1 = 0$. Assim, $r \equiv 0 \pmod 4$.

Agora consideramos o fator 3^2 de $q-1$ e determinamos $r = r_0 + 3r_1 \pmod 9$. Como $62^{60} = 48$ deduzimos de $\alpha^{(q-1)/p_2} = a^{(q-1)/p_2}$ que $r_0 = 1$. Para determinar r_1, consideramos $62 \cdot 2^{-1} = 31$ e $31^{180/9} = 1$, portanto $r_1 = 0$. No total, $r \equiv 1 \pmod 9$.

Para o fator 5 de $q - 1$ temos que determinar $r \equiv r_0 \pmod 5$. Como $62^{36} = 1$, concluímos que $r_0 = 0$ e $r \equiv 0 \pmod 5$.

Finalmente, resolvendo as três congruências $r \equiv 0 \pmod 4$, $r \equiv 1 \pmod 9$ e $r \equiv 0 \pmod 5$ simultaneamente obtemos $r \equiv 100 \pmod{180}$, portanto $r = 100$.

O criptossistema de chave pública baseado no problema da *"mochila"* (*knapsack*) explora a dificuldade computacional de selecionar, dentre uma coleção de objetos com pesos distintos, um subconjunto cuja soma seja exatamente igual a um valor pré-determinado. Em termos matemáticos, dado um vetor $y = (y_1, \ldots, y_n)$, escolhe-se um vetor binário $a = (a_1, \ldots, a_n)$, onde $a_i = 1$ indica que o item y_i foi escolhido e $a_i = 0$ indica que não foi. O problema consiste em determinar a a partir de $K = y \cdot a$, produto interno dos vetores y e a. Quando apenas K e y são conhecidos, encontrar a é, em geral, difícil, pertencendo à classe dos problemas NP-completos. Contudo, existem casos especiais simples, como quando o vetor y é *supercrescente*, isto é, $y_i > y_1 + \cdots + y_{i-1}$ para $i \geq 2$. Nessa situação, a solução pode ser determinada em tempo linear por um algoritmo direto.

A ideia principal do criptossistema é transformar uma "mochila" simples, facilmente solucionável, em uma "mochila"complexa, cujo problema inverso seja computacionalmente difícil, criando assim a chamada *"mochila com armadilha"* (*trapdoor knapsack*). Para isso escolhem-se inteiros grandes k e q com $mdc(k, q) = 1$, e define-se o novo vetor público $z = (z_1, \ldots, z_n)$ por

$$z_i = ky_i \pmod{q}.$$

O vetor z é divulgado como chave pública, enquanto y, k e q formam a chave privada. Para cifrar uma mensagem binária $a = (a_1, \ldots, a_n)$, o emissor calcula

$$L = a_1 z_1 + \cdots + a_n z_n,$$

e transmite L. Já o receptor autorizado computa $k^{-1}L \pmod{q}$, recuperando

$$K = y \cdot a,$$

o que reduz o problema a uma "mochila"supercrescente fácil de resolver, permitindo a obtenção da mensagem original. Embora tenham sido propostos tamanhos grandes para n e condições adicionais sobre q, y e k, ataques desenvolvidos a partir de 1982 mostraram vulnerabilidades significativas, colocando em dúvida a segurança do método de Merkle e Hellman. Ainda assim, o sistema permanece como um exemplo instrutivo de como um problema difícil pode ser transformado em uma função de mão única com atalho utilizável em criptografia.

Definição 123 (Criptografia da mochila com armadilha)**.** *Para realizar a transmissão de uma mensagem em texto original, deve-se primeiramente estabelecer um mapeamento entre os caracteres do alfabeto e os elementos de $\mathbb{Z}_2$, o que equivale a codificar cada letra como uma sequência binária de comprimento t. Em seguida, selecionam-se os parâmetros q e n de modo que t seja um divisor de n ($t \mid n$), e define-se um vetor $y \in \mathbb{Z}_q^n$ com a propriedade de que cada componente y_i seja estritamente maior que a soma de todos os componentes anteriores ($y_i > y_1 + \cdots + y_{i-1}$) para todo índice i entre 2 e n. Adicionalmente, escolhe-se um inteiro k que seja coprimo com q ($mdc(k, q) = 1$). Os processos de cifragem e decifragem são então realizados conforme descrito abaixo:*

Processo de cifragem: *Calcula-se o vetor $x := ky$. A mensagem original é segmentada em blocos, cada um contendo n/t caracteres. Estes blocos são subsequentemente convertidos em sequências binárias de comprimento n. Cada um desses blocos binários é então multiplicado pelo vetor x, realizando-se a operação módulo q.*

Processo de Decifragem: *A sequência cifrada recebida é multiplicada pelo inverso multiplicativo de k em $\mathbb{Z}_q$ (denotado por k^{-1}, cuja existência é garantida pela condição $mdc(k, q) = 1$). Como resultado, cada elemento da sequência decifrada corresponde diretamente a uma única sequência binária de comprimento n. Esta sequência binária é, por fim, reinterpretada como um conjunto de n/t grupos, onde cada grupo é composto por t bits que representam o caractere original.*

Exemplo 124. *Seja $t = 5$, ou seja, substituímos as letras do alfabeto por 5-tuplas de elementos em $\mathbb{Z}_2$.*

A	*B*	*C*	*D*	*E*	*F*	*G*	*H*	*I*	*J*	*K*
00000	*00001*	*00010*	*00011*	*00100*	*00101*	*00110*	*00111*	*01000*	*01001*	*01010*
L	*M*	*N*	*O*	*P*	*Q*	*R*	*S*	*T*	*U*	*V*
01011	*01100*	*01101*	*01110*	*01111*	*10000*	*10001*	*10010*	*10011*	*10100*	*10101*
W	*X*	*Y*	*Z*	*,*	*.*	*!*	*?*	*”*	*“espaço”*	
10110	*10111*	*11000*	*11001*	*11010*	*11011*	*11100*	*11101*	*11110*	*11111*	

Agora escolhemos $n = 10$, $q = 19999$, e um vetor

$$y = (4, 10, 64, 101, 200, 400, 800, 1980, 4000, 9000) \in (\mathbb{Z}_{19999})^{10}.$$

Em y cada coordenada é maior que a soma das coordenadas anteriores se considerarmos os elementos como inteiros, portanto y é supercrescente, formando uma mochila simples. Formamos um novo vetor

$$x = 200y = (800, 2000, 12800, 201, 2, 4, 8, 16019, 40, 90) \in (\mathbb{Z}_{19999})^{10}.$$

Se desejamos cifrar a palavra HUMBERTO subdividimos a mensagem em grupos de dois: HU MB ER TO e codificamos:

$$\begin{aligned} HU &= 0011110100 \\ &\to (0,0,1,1,1,1,0,1,0,0) \cdot (800, 2000, 12800, 201, 2, 4, 8, 16019, 40, 90) \\ &= 9027. \end{aligned}$$

Também:

$$\begin{aligned} MB &\to 14890, \\ ER &\to 12894, \\ TO &\to 17070 \end{aligned}$$

Assim, o texto plano cifrado é

$$(9027, 14890, 12894, 17070) \in (\mathbb{Z}_{19999})^4.$$

Se o receptor obtém a mensagem

$$w = (18069, 17020) \in (\mathbb{Z}_{19999})^2,$$

então a deciframos da seguinte forma. Avaliamos $200^{-1}w = 100w \in (\mathbb{Z}_{19999})^2$*. Para a primeira componente, digamos, obtemos* $6990 = \mathbf{a} \cdot y = 4a_1 + 10a_2 + \cdots + 9000a_{10}$*, onde* $\mathbf{a} = (a_1, \ldots, a_{10})$*, com* $a_i \in \{0, 1\}$*, representa a primeira parte original da mensagem em forma numérica.*

6990 só pode ser representado na forma

$$1 \cdot 4000 + 1 \cdot 1980 + 1 \cdot 800 + 1 \cdot 200 + 1 \cdot 10 = 10a_2 + 200a_5 + 800a_7 + 1980a_8 + 4000a_9,$$

se subtrairmos de 6990 os números 4000, 1980, etc. Assim obtemos

$$\mathbf{a} = (0, 1, 0, 0, 1, 0, 1, 1, 1, 0) \leftrightarrow 01001\,01110 \rightarrow JO.$$

O receptor continua decifrando e lê a mensagem SE, o que dá o texto plano JOSE.

O criptossistema de mochila foi inicialmente proposto como um método promissor para comunicação segura, pois permitia que cada usuário publicasse uma chave aberta — o vetor x — enquanto mantinha em segredo os parâmetros k e q, responsáveis por transformar uma mochila simples em uma mochila complexa. Isso tornava possível que grandes organizações, como grupos bancários com muitos participantes, compartilhassem chaves públicas sem a necessidade de estabelecer um grande número de acordos secretos individuais. Contudo, verificou-se que o sistema é vulnerável a ataques criptanalíticos eficientes, mesmo quando se tenta reforçá-lo por meio de iterações sucessivas (múltiplas transformações da mochila).

De modo mais amplo, esta seção enfatiza que desafios importantes na criptografia moderna incluem estabelecer a segurança do RSA (ou provar sua equivalência ao problema de fatoração) e desenvolver novos criptossistemas públicos realmente seguros. Embora os sistemas de chave pública eliminem a necessidade de distribuição de chaves secretas, eles exigem mecanismos adicionais de autenticação, já que qualquer pessoa pode gerar mensagens cifradas com uma chave pública. Sistemas como o RSA destacam-se nesse contexto, pois permitem combinar sigilo e autenticação por meio de assinaturas digitais que apenas o remetente legítimo pode produzir.

4.2 Teoria de Códigos

4.2.1 Introdução à Teoria de Códigos

A **Teoria de Códigos** emerge como um campo essencial da matemática e da engenharia focado em solucionar um desafio fundamental dos sistemas modernos de comunicação: a transmissão de dados sem erros através de canais ruidosos. Dispositivos eletrônicos e meios de comunicação, como redes de computadores e transmissões via satélite, são inerentemente suscetíveis a falhas que introduzem ruído, alterando a integridade da

informação original. O objetivo central desta teoria é equipar o sinal com redundância controlada, permitindo que o receptor não apenas detecte a presença de erros, mas também os corrija, recuperando a mensagem original com alta fidelidade.

Este processo é formalizado pelo modelo de comunicação: uma **mensagem original a** é submetida a um **codificador** f, que gera uma **palavra-código c** de maior comprimento. Essa palavra **c** atravessa um **canal ruidoso**, onde sofre a adição de um **vetor de erro e**, resultando no **vetor recebido** $\mathbf{y} = \mathbf{c} + \mathbf{e}$. O **decodificador** g no lado receptor utiliza as propriedades do código para inferir a palavra-código original e, consequentemente, a mensagem $\overline{\mathbf{a}}$.

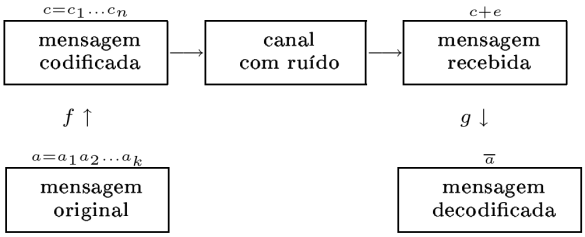

Figura 4 – Diagrama Teoria de Códigos . Presente no livro (PANARIO, 2007)

O estudo dessas técnicas é estruturado por sua complexidade e capacidade algébrica, iniciando-se pelos **Códigos Lineares**, que se baseiam nos princípios da Álgebra Linear, sendo definidos como subespaços vetoriais. Em seguida, exploramos os **Códigos Cíclicos**, uma subclasse poderosa dos códigos lineares, cuja estrutura é elegantemente representada por ideais em anéis de polinômios. Esta representação simplifica o projeto de codificadores e decodificadores. Por fim, esta introdução culmina com uma menção breve, mas fundamental, aos **Códigos BCH** (Bose-Chaudhuri-Hocquenghem). Estes códigos representam um avanço significativo, sendo notáveis por sua alta capacidade de correção de múltiplos erros e por constituírem uma generalização dos Códigos de Hamming.

Definição 125 (Código Linear)**.** *Seja $\mathbb{F}_q$ um corpo finito com q elementos (usualmente $q = 2$). Um subconjunto C do espaço vetorial $\mathbb{F}_q^n$ é denominado um* ***código linear*** *(n, k) se C for um subespaço vetorial de dimensão k sobre $\mathbb{F}_q$.*

Neste contexto, n representa o comprimento do código (o número de símbolos na palavra-código) e k representa a dimensão do código (o número de símbolos da mensagem original). A taxa de informação do código é dada por $R = k/n$.

Em um Código Linear de dimensão (n, k), a **Matriz de Paridade** (H) é uma matriz de dimensão $(n - k) \times n$ cuja função principal é a verificação de palavras-código.

Um vetor $c = (c_1, c_2, \ldots, c_n)$ é uma palavra-código válida se, e somente se, ele pertencer ao código C, que é o subespaço vetorial de dimensão k. O código C é definido pelo sistema de equações lineares homogêneas geradas pela matriz de paridade H:

$$Hc^T = \mathbf{0}$$

Exemplo 126. *Um dos códigos lineares mais básicos e didáticos é o código verificador de paridade simples. Este é um código $(n, n-1)$ que possui a função primária de detectar um único erro na transmissão.*

*A cada $k = n-1$ bits de informação $\mathbf{a} = (a_1, a_2, \ldots, a_{n-1})$, um n-ésimo bit, chamado de **bit de paridade** a_n, é adicionado. O bit de paridade é escolhido de tal forma que a soma (em $\mathbb{F}_2$) de todos os bits na palavra-código resultante $\mathbf{c} = (a_1, a_2, \ldots, a_n)$ seja zero:*

$$\sum_{i=1}^{n} a_i = 0 \quad (em\ \mathbb{F}_2)$$

Se o receptor receber um vetor $\mathbf{y}$ cuja soma dos bits é 1, ele sabe que houve um número ímpar de erros. Se a soma for 0, ele assume que não houve erro (embora não possa detectar um número par de erros, como dois).

Por exemplo, para $n = 4$ (um código $(4, 3)$):

- *A mensagem $\mathbf{a} = (1, 1, 0)$ é codificada como $\mathbf{c} = (1, 1, 0, 0)$, pois $1 + 1 + 0 + 0 = 0$.*
- *A mensagem $\mathbf{a} = (1, 0, 0)$ é codificada como $\mathbf{c} = (1, 0, 0, 1)$, pois $1 + 0 + 0 + 1 = 0$.*

Exemplo 127 (Código de Repetição)**.** *O **Código de Repetição** é um exemplo clássico de código linear do tipo $(n, 1)$, cuja principal característica é repetir várias vezes o mesmo símbolo de mensagem. A ideia central é aumentar significativamente a redundância, tornando o código robusto contra erros introduzidos pelo canal.*

Considere uma mensagem constituída por um único símbolo

$$a_1 \in \mathbb{F}_q.$$

O processo de codificação consiste em repetir esse símbolo nas n posições da palavra-código, obtendo

$$f(a_1) = (a_1, a_1, \ldots, a_1) \in \mathbb{F}_q^n.$$

Este código é linear e pode ser descrito por uma matriz geradora

$$G = (1\,1\,\cdots\,1),$$

e por uma matriz de paridade da forma

$$H = \begin{pmatrix} -1 & 1 & 0 & \cdots & 0 \\ -1 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -1 & 0 & 0 & \cdots & 1 \end{pmatrix},$$

onde I_{n-1} denota a matriz identidade de ordem $n-1$. O sistema

$$Hc^T = 0$$

gera as equações

$$-c_1 + c_2 = 0, \quad -c_1 + c_3 = 0, \quad \ldots, \quad -c_1 + c_n = 0,$$

das quais concluímos que toda palavra-código válida deve satisfazer

$$c_1 = c_2 = \cdots = c_n.$$

No contexto binário ($q = 2$), a decodificação é feita por decisão por maioria*: dado um vetor recebido $\mathbf{y}$, o decodificador escolhe como mensagem estimada o símbolo que ocorre com maior frequência nas coordenadas de $\mathbf{y}$.*

Por exemplo, em um código de repetição $(5, 1)$:

$$a_1 = 1 \quad \longmapsto \quad f(a_1) = (1, 1, 1, 1, 1).$$

Se o vetor recebido for

$$\mathbf{y} = (1, 0, 1, 1, 0),$$

então há três ocorrências de 1 e duas de 0, de modo que o decodificador decide por

$$\overline{a}_1 = 1.$$

Esse método de decodificação permite corrigir erros desde que mais da metade das posições recebidas permaneça correta. Assim, o código de repetição possui grande capacidade de detecção e correção, embora à custa de uma taxa de transmissão muito baixa.

No código de repetição, quando há apenas um símbolo, usa-se n ímpar, para evitar a igualdade de ocorrências entre as coordenadas (e assim garantir a decodificação por maioria).

O código de repetição possui uma elevada capacidade de detecção de erros. Ele consegue detectar até $n-1$ erros, pois, sempre que ao menos dois símbolos recebidos forem diferentes, concluímos que ocorreu alguma alteração durante a transmissão. Entretanto, caso todos os n símbolos recebidos sejam idênticos, não é possível determinar se houve ou não erro, já que todos os símbolos podem ter sido modificados da mesma forma. Assim, o código não detecta erros quando exatamente n símbolos são alterados.

Por outro lado, o código de repetição permite corrigir erros sempre que sua quantidade não exceder

$$\left\lfloor \frac{n-1}{2} \right\rfloor.$$

Definição 128 (Distância de Hamming)**.** *Sejam x e y vetores em $\mathbb{F}_q^n$. A distância de Hamming $d(x, y)$ é o número de coordenadas nas quais x e y diferem entre si.*

A distância de Hamming é uma métrica em $\mathbb{F}_q^n$, isto é, satisfaz as seguintes propriedades: para quaisquer $x, y, z \in \mathbb{F}_q^n$, temos

1. $d(x, y) \geq 0$,
2. $d(x, y) = 0$ *se e somente se* $x = y$,
3. $d(x, y) = d(y, x)$,
4. $d(x, z) \leq d(x, y) + d(y, z)$.

Um conceito relacionado à distância de Hamming é o peso de Hamming. O peso de Hamming é o número de coordenadas não nulas de x e é denotado por $w(x)$.

Definição 129. *Se $\mathbf{c}$ é uma palavra-código e $\mathbf{y}$ é a palavra recebida, então o* erro *é dado pela diferença*

$$\mathbf{e} = \mathbf{y} - \mathbf{c}.$$

Definição 130. *Seja t um inteiro positivo. Dizemos que um código $C \subseteq \mathbb{F}_q^n$* corrige *$t$ erros quando, para toda palavra recebida $\mathbf{y} \in \mathbb{F}_q^n$, existe no máximo uma palavra-código $\mathbf{c} \in C$ tal que*

$$d(\mathbf{c}, \mathbf{y}) \leq t.$$

Em outras palavras, se a bola fechada de raio t centrada em $\mathbf{y}$ contém duas palavras-código distintas, então o código não é capaz de corrigir t erros.

Definição 131. *A* distância mínima *de um código C é definida por*

$$d_C = \min_{\substack{\mathbf{u}, \mathbf{v} \in C \\ \mathbf{u} \neq \mathbf{v}}} d(\mathbf{u}, \mathbf{v}) = \min_{\substack{\mathbf{c} \in C \\ \mathbf{c} \neq \mathbf{0}}} w(\mathbf{c}),$$

onde $w(\mathbf{c})$ denota o peso de Hamming da palavra $\mathbf{c}$, que é o numero de coordenadas não nulas de c.

Teorema 132. *Um código C com distância mínima d_C pode corrigir até t erros, se $d_C \geq 2t + 1$.*

Demonstração. A bola fechada $B_t(x)$ de raio t e centro $x \in \mathbb{F}_q^n$ é formada por todos os vetores $y \in \mathbb{F}_q^n$ tais que $d(x, y) \leq t$. A regra de decodificação por verossimilhança garante que cada palavra recebida com no máximo t erros deve estar numa bola com centro na palavra transmitida e raio t. Suponhamos que $u \in B_t(x) \cap B_t(y)$ onde $x, y \in C$. Então $d(x, y) \leq d(x, u) + d(u, y) \leq 2t$, o que é uma contradição, pois $d_C \geq 2t + 1$. □

Para calcular d_C de forma alternativa basta usarmos o seguinte resultado

Teorema 133. *Considere um código linear C cuja matriz de paridade é H. Então o código possui distância mínima $d_C \geq s+1$ se, e somente se, quaisquer s colunas de H forem linearmente independentes.*

Demonstração. Primeiro, suponha que H possua s colunas linearmente dependentes, digamos

$$D_{i_1}, D_{i_2}, \ldots, D_{i_s}.$$

Então existem escalares

$$\alpha_1, \alpha_2, \ldots, \alpha_s \in \mathbb{F}_q,$$

não todos iguais a zero, tais que

$$\sum_{j=1}^{s} \alpha_j D_{i_j} = 0.$$

Construímos então um vetor $c \in \mathbb{F}_q^n$ definindo

$$c_{i_j} = \alpha_j \quad \text{para } j = 1, 2, \ldots, s, \qquad c_\ell = 0 \text{ para as demais coordenadas.}$$

Como

$$Hc^T = \sum_{j=1}^{s} \alpha_j D_{i_j} = 0,$$

segue que $c \in C$. Além disso, $c \neq 0$ (pois algum α_j é não nulo) e seu peso satisfaz

$$w(c) \leq s.$$

Portanto, existe uma palavra-código não nula com peso no máximo s, o que implica que a distância mínima não pode ser maior que s, isto é,

$$d_C \leq s.$$

Agora suponha o contrário: que qualquer conjunto de s colunas de H é linearmente independente.

Mostremos que, com essa condição, não existe palavra-código não nula com peso menor ou igual a s. Com efeito, suponha que exista $c \in C$, $c \neq 0$, tal que $w(c) \leq s$. Consideremos as posições em que c é não nulo,

$$i_1, i_2, \ldots, i_r, \qquad r = w(c) \leq s.$$

Como $c \in C$, temos

$$Hc^T = \sum_{j=1}^{r} c_{i_j} D_{i_j} = 0,$$

o que fornece uma combinação linear nula envolvendo apenas as colunas $D_{i_1}, \ldots, D_{i_r}$. Isso contraria a hipótese de independência linear de qualquer conjunto de até s colunas de H.

Logo, tal vetor c não pode existir, e a única palavra com peso $\leq s$ é a palavra nula. Consequentemente,

$$d_C \geq s + 1.$$

Isto estabelece a equivalência desejada. □

Exemplo 134. *Considere a matriz de verificação de paridade H sobre o corpo binário $\mathbb{F}_2$:*

$$H = \begin{pmatrix} 1 & 1 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 1 \end{pmatrix}.$$

Observa-se que quaisquer duas colunas de H são linearmente independentes. Pelo Teorema 133, isso garante que a distância mínima d_C do código C definido por H satisfaz $d_C \geq 3$.

No entanto, note que a coluna 4 é igual à soma (em $\mathbb{F}_2$) da coluna 3 e da coluna 5:

$$\textit{Coluna 4} = (1,1,1)^\top, \quad \textit{Coluna 3} = (0,1,1)^\top, \quad \textit{Coluna 5} = (1,0,0)^\top,$$

e verifica-se que:

$$(0,1,1)^\top + (1,0,0)^\top = (1,1,1)^\top.$$

Isso significa que três colunas de H são linearmente dependentes. Portanto, a distância mínima d_C não pode ser 4 ou mais; conclui-se que $d_C = 3$.

Pelo Teorema 132, um código com distância mínima $d_C = 3$ satisfaz $3 \geq 2t + 1$, ou seja, $t \leq 1$. Logo, o código pode corrigir até 1 erro.

Vamos agora abordar o problema da decodificação de uma palavra recebida y. Uma abordagem direta consiste em calcular a distância entre y e todas as palavras-código possíveis, selecionando para correção aquela que estiver mais próxima.

Entretanto, essa estratégia se torna computacionalmente inviável quando a dimensão k do código é grande, uma vez que exige a realização de q^k cálculos. Diante dessa limitação, apresentaremos a seguir uma técnica de decodificação mais eficiente.

Tomemos um código linear $C(n,k)$ definido sobre o corpo finito $\mathbb{F}_q$. Sejam

$$0 = c^{(1)}, c^{(2)}, \ldots, c^{(n')}$$

as palavras-código distintas. Note que C constitui um subespaço vetorial de $\mathbb{F}_q^n$. O espaço quociente $\mathbb{F}_q^n/C$ é composto pelas classes laterais da forma $a + C = \{a + c : c \in C\}$, com $a \in \mathbb{F}_q^n$.

Algumas propriedades importantes:

- Cada classe lateral possui exatamente $|C| = q^k$ elementos.
- O número total de classes laterais é q^{n-k}.

Dessa forma, podemos decompor o espaço $\mathbb{F}_q^n$ na união disjunta:

$$\mathbb{F}_q^n = (a^{(0)} + C)\dot{\cup}(a^{(1)} + C)\dot{\cup}\cdots\dot{\cup}(a^{(s)} + C),$$

onde $a^{(0)} = 0$ e $s = q^{n-k} - 1$.

Suponha que uma palavra $y \in \mathbb{F}_q^n$ seja recebida, de modo que $y \in a^{(i)} + C$ para algum i, ou mais precisamente, $y = a^{(i)} + c^{(j)}$, com $0 \leq i \leq s$ e $1 \leq j \leq q^k$. Se a palavra originalmente transmitida foi c, então o vetor erro é dado por $e = y - c = a^{(i)} + c^{(j)} - c$. Devido ao fato de C ser um espaço vetorial (e, portanto, fechado sob adição), o erro e pertence à mesma classe lateral que a palavra recebida y.

Definição 135. *Seja H uma matriz de verificação de paridade para um código linear $C(n, k)$ sobre $\mathbb{F}_q$, e seja $y \in \mathbb{F}_q^n$ um vetor qualquer. O vetor $S(y) = Hy^\top \in \mathbb{F}_q^{n-k}$ é denominado a síndrome de y.*

Teorema 136. *Para quaisquer vetores $y, z \in \mathbb{F}_q^n$, valem as seguintes propriedades:*

1. *$S(y) = 0$ se e somente se $y \in C$.*
2. *$S(y) = S(z)$ se e somente se y e z pertencem à mesma classe lateral de C, ou seja, $y + C = z + C$.*

Demonstração. 1. Pela definição, $S(y) = Hy^\top$. Se $y \in C$, então, pela definição de matriz de verificação de paridade, $Hy^\top = 0$, logo $S(y) = 0$. Reciprocamente, se $S(y) = 0$, então $Hy^\top = 0$, o que significa que y pertence ao espaço nulo de H, que é exatamente o código C. Portanto, $y \in C$.

2. Observe que $S(y) = S(z)$ se e somente se $Hy^\top = Hz^\top$. Isto é equivalente a $H(y^\top - z^\top) = 0$, ou seja, $H(y - z)^\top = 0$. Pela definição do código, isto significa que $y - z \in C$. Dois vetores pertencem à mesma classe lateral módulo C se, e somente se, a sua diferença é um elemento do código. Assim, $y - z \in C$ é equivalente a $y + C = z + C$.

□

Definição 137. *Seja $C(n,k)$ um código linear sobre $\mathbb{F}_q$. Em cada classe lateral $a+C \subseteq \mathbb{F}_q^n$, chamamos de* líder da classe *um vetor de peso mínimo dentro dessa classe. Caso haja mais de um vetor com peso mínimo, escolhe-se qualquer um deles.*

Podemos organizar os elementos de $\mathbb{F}_q^n$ segundo suas classes laterais, colocando no topo de cada classe o seu líder. Assim, cada vetor recebido pertence a uma dessas classes, e se o erro estiver na classe cujo líder é $a^{(i)}$, então o erro é precisamente esse líder. Nesse caso, a mensagem transmitida pode ser recuperada por

$$x = y - e = a^{(i)} + c^{(j)} - a^{(i)} = c^{(j)}.$$

Determinar os líderes das classes laterais fornece um método eficiente de decodificação. A síndrome permite identificar a classe lateral correspondente ao vetor recebido, e então basta utilizar o líder dessa classe, que é o vetor de menor peso, o que está diretamente relacionado à distância mínima do código.

Seja $C(4,2)$ um código binário linear com matriz geradora:

$$G = \begin{pmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 \end{pmatrix}$$

e matriz de verificação de paridade:

$$H = \begin{pmatrix} 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix}.$$

Existem 4 mensagens distintas, 4 palavras-código e 4 classes laterais. A tabela a seguir organiza os elementos de $\mathbb{F}_2^4$ particionados em classes de acordo com este código. A distribuição desta tabela será explicada em detalhes.

linha mensagem	00	01	10	11	**Síndrome**
palavras-código	0000	0110	1011	1101	$\begin{pmatrix}0\\0\end{pmatrix}$
classes laterais restantes	0001	0111	1010	1100	$\begin{pmatrix}0\\1\end{pmatrix}$
classes laterais restantes	0010	0100	1001	1111	$\begin{pmatrix}1\\0\end{pmatrix}$
classes laterais restantes	1000	1110	0011	0101	$\begin{pmatrix}1\\1\end{pmatrix}$

As palavras-código são calculadas usando $\mathbf{a}G = \mathbf{c}$ ou $H\mathbf{c}^T = \mathbf{0}$. Os elementos das demais classes laterais são obtidos somando um líder de classe lateral a cada palavra-código.

Observa-se que a tabela completa pode ser construída conhecendo-se apenas os líderes das classes laterais e suas síndromes correspondentes.

Para determinar os líderes das classes laterais, consideramos os vetores sequencialmente, verificando se cada síndrome é nova ou já foi encontrada. Se for nova, armazenamos temporariamente esse elemento como líder da classe lateral. Caso já exista um líder para aquela síndrome, comparamos os pesos de Hamming, atualizando o líder sempre que um vetor de peso menor for encontrado. Ao final do processo, todos os líderes das classes laterais estarão determinados. Existem diversas estratégias para otimizar este algoritmo, como por exemplo, considerar os vetores ordenados por peso crescente.

Se a palavra recebida é $\mathbf{y} = 0101$, então $S(\mathbf{y}) = \binom{1}{1}$. Pela tabela, tomamos o líder da classe lateral correspondente (vetor erro) como $\mathbf{e} = 1000$. A mensagem transmitida original é recuperada subtraindo o erro da palavra recebida:

$$\mathbf{x} = \mathbf{y} - \mathbf{e} = 0101 - 1000 = 0101 + 1000 = 1101.$$

Portanto, a palavra-código transmitida foi 1101, que corresponde à mensagem original 11.

Teorema 138. *Considere um código linear binário $C(n, k)$ com matriz de paridade H. Para uma palavra recebida $\mathbf{y} \in \mathbb{F}_2^n$, a síndrome $S(\mathbf{y})$ é igual à soma das colunas de H correspondentes às posições onde ocorreram erros.*

Demonstração. Seja $\mathbf{y} = \mathbf{x} + \mathbf{e}$, onde $\mathbf{x}$ é a palavra-código enviada e $\mathbf{e}$ é o vetor de erros. Como $\mathbf{x} \in C$, sabemos que $S(\mathbf{x}) = \mathbf{0}$. Logo, pela linearidade:

$$S(\mathbf{y}) = S(\mathbf{x} + \mathbf{e}) = S(\mathbf{x}) + S(\mathbf{e}) = H(\mathbf{x} + \mathbf{e})^T = H\mathbf{e}^T.$$

Supondo que os erros ocorram nas posições $i_1, i_2, \ldots, i_t$, o vetor de erro tem 1 nessas posições e 0 nas demais. Assim, o produto $H\mathbf{e}^T$ seleciona e soma as colunas correspondentes:

$$S(\mathbf{y}) = H\mathbf{e}^T = H_{i_1} + H_{i_2} + \cdots + H_{i_t},$$

onde H_j denota a j-ésima coluna de H. □

Se o vetor de erro possui somente um erro, a síndrome identifica diretamente a coluna de H correspondente, permitindo a correção. Se houver mais de um erro, a síndrome passa a ser a soma de várias colunas, e o decodificador pode não conseguir recuperar a mensagem. Assim, este método garante correção de um único erro.

4.2.2 Códigos Cíclicos e BCH

Um código cíclico é um código linear tal que o deslocamento cíclico dos símbolos em uma palavra-código produz outra palavra-código válida. Esta propriedade de fechamento

sob operações de deslocamento cíclico confere aos códigos cíclicos uma estrutura algébrica rica que facilita sua implementação e análise.

Os códigos cíclicos formam uma subclasse importante dos códigos corretores de erro, amplamente utilizados em sistemas de comunicação digital, armazenamento de dados e aplicações que requerem transmissão confiável de informação através de canais ruidosos.

Definição 139. *Um código linear $C(n,k)$ definido sobre o corpo finito $\mathbb{F}_q$ é denominado* ***código cíclico*** *quando apresenta a seguinte propriedade: se uma sequência $(c_0, c_1, \ldots, c_{n-1})$ é uma palavra-código válida em C, então a sequência obtida pelo deslocamento circular desses elementos $(c_{n-1}, c_0, c_1, \ldots, c_{n-2})$ também pertence ao código C.*

Exemplo 140. *Considere o código de Hamming C_3 sobre $\mathbb{F}_2$ com matriz de verificação de paridade:*

$$H = \begin{pmatrix} 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 & 1 \end{pmatrix}$$

Tomemos o vetor $c = (0,1,0,1,1,0,0)$. Este vetor é uma palavra-código válida, uma vez que:

$$Hc^T = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

Aplicando uma operação de deslocamento cíclico ao vetor c, obtemos $\overline{c} = (0,0,1,0,1,1,0)$. Verificamos que este novo vetor também satisfaz a condição de paridade:

$$H\overline{c}^T = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

confirmando que $\overline{c}$ é igualmente uma palavra-código válida. De fato, este código possui a propriedade cíclica. Para demonstrar esta característica de forma geral, expressamos as palavras-código na forma:

$$(r+s+t,\ s+t+u,\ r+s+u,\ r,\ s,\ t,\ u)$$

onde $r, s, t, u \in \mathbb{F}_2$. Como podemos verificar que:

$$H \begin{pmatrix} u & r+s+t & s+t+u & r+s+u & r & s & t \end{pmatrix}^T = 0,$$

concluímos que qualquer deslocamento cíclico de uma palavra-código resulta em outra palavra-código válida, estabelecendo assim o caráter cíclico do código.

Identificamos $\mathbb{F}_q^n$ com o anel $\mathbb{F}_q[x]/\langle x^n-1\rangle$, onde o vetor $c = (c_0, \ldots, c_{n-1})$ corresponde ao polinômio $c(x) = c_0 + c_1x + \cdots + c_{n-1}x^{n-1}$.

O deslocamento cíclico equivale à multiplicação por x:

$$x \cdot c(x) = c_{n-1} + c_0x + \cdots + c_{n-2}x^{n-1},$$

pois $x^n \equiv 1 \pmod{x^n - 1}$.

Um código $C \subseteq \mathbb{F}_q^n$ é cíclico se e somente se $C(x)$ é um ideal em $\mathbb{F}_q[x]/\langle x^n - 1\rangle$.

Teorema 141. *Um código linear $C(n, k)$ sobre $\mathbb{F}_q$ é cíclico se, e somente se, C é um ideal de $\mathbb{F}_q[x]/\langle x^n - 1\rangle$.*

Demonstração. Dividimos a demonstração em duas partes:

($\Rightarrow$) Suponha que C é um código cíclico. Seja $c(x) \in C$. Como C é cíclico, temos que $xc(x), x^2c(x), x^3c(x), \ldots$ também pertencem a C. Seja $a(x) = \sum_i a_ix^i \in \mathbb{F}_q[x]/\langle x^n - 1\rangle$. Então:

$$a(x)c(x) = \sum_i a_i(x^ic(x))$$

Como cada termo $x^ic(x)$ pertence a C (por ser combinação de deslocamentos cíclicos) e C é um subespaço vetorial sobre $\mathbb{F}_q$, segue que $a(x)c(x) \in C$. Portanto, C é um ideal.

($\Leftarrow$) Reciprocamente, se C é um ideal de $\mathbb{F}_q[x]/\langle x^n - 1\rangle$ e $c(x) = \sum_{i=0}^{n-1} c_ix^i$ é uma palavra-código, então $xc(x)$ também pertence a C, ou seja, é uma palavra-código. Logo, o código C é cíclico. □

Observemos que, como todo ideal não nulo de $\mathbb{F}_q[x]/\langle x^n - 1\rangle$ é necessariamente principal e possui um gerador mônico, segue que cada código cíclico pode ser descrito por meio de um único polinômio gerador.

Definição 142. *Seja $C = \langle g\rangle$ um código cíclico. Dizemos que g é o polinômio gerador de C e $h = (x^n - 1)/g$ é o polinômio verificador de C.*

Teorema 143. *Seja C um ideal não trivial no anel quociente $\mathbb{F}_q[x]/\langle x^n - 1\rangle$, ou seja, C é um código cíclico de comprimento n. Então valem as seguintes propriedades:*

1. *O código C possui um único polinômio gerador mônico g que tem o menor grau entre todos os polinômios não nulos em C.*
2. *O polinômio gerador g divide $x^n - 1$ em $\mathbb{F}_q[x]$.*

3. *No anel* $\mathbb{F}_q[x]$*, todo elemento* $c \in C$ *pode ser expresso de maneira única como* $c = fg$*, onde* $\deg(f) < n - r$ *e* $\deg(g) = r$*. Adicionalmente, a dimensão do código* C *é* $n - r$*.*

4. *Se* $g(x) = g_0 + g_1 x + \cdots + g_r x^r$*, então* C *é gerado como subespaço vetorial de* $\mathbb{F}_q^n$ *pelas linhas da matriz geradora:*

$$G = \begin{pmatrix} g_0 & g_1 & \cdots & g_r & 0 & 0 & \cdots & 0 \\ 0 & g_0 & g_1 & \cdots & g_r & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & g_0 & g_1 & \cdots & \cdots & g_r \end{pmatrix}$$

que corresponde aos polinômios $g(x), xg(x), x^2 g(x), \ldots, x^{n-r-1} g(x)$*.*

Demonstração. A existência e unicidade do polinômio gerador mônico de grau mínimo decorre do fato de que $\mathbb{F}_q[x]$ é um domínio de ideais principais. No anel quociente $\mathbb{F}_q[x]/(x^n - 1)$, todo ideal não nulo é gerado por um único polinômio mônico de grau mínimo.

Para demonstrar que g divide $x^n - 1$, consideremos a divisão polinomial em $\mathbb{F}_q[x]$:

$$x^n - 1 = h(x)g(x) + r(x),$$

onde $\deg(r(x)) < \deg(g(x))$. No anel quociente $\mathbb{F}_q[x]/(x^n - 1)$, esta equação se torna:

$$r(x) = -h(x)g(x) \mod (x^n - 1).$$

Como C é um ideal e $g(x) \in C$, segue que $r(x) \in C$. Porém, $\deg(r(x)) < \deg(g(x))$ e $g(x)$ é o polinômio de grau mínimo em C, obrigando $r(x) = 0$. Portanto, $g(x)$ divide $x^n - 1$.

Seja $c \in C$ com $\deg(c) < n$. Pelo item (2), existe $h(x) \in \mathbb{F}_q[x]$ tal que $x^n - 1 = g(x)h(x)$. Como $c \in C = (g(x))$, existe $f(x) \in \mathbb{F}_q[x]$ tal que:

$$c(x) = f(x)g(x) + \ell(x)(x^n - 1) = [f(x) + \ell(x)h(x)]g(x).$$

Definindo $\tilde{f}(x) = f(x) + \ell(x)h(x)$, temos $c(x) = \tilde{f}(x)g(x)$. Podemos escolher $\tilde{f}(x)$ tal que $\deg(\tilde{f}) < n - r$, onde $r = \deg(g)$, através do algoritmo da divisão. A unicidade segue da condição sobre o grau.

Os polinômios $g(x), xg(x), x^2 g(x), \ldots, x^{n-r-1} g(x)$ formam uma base para C, portanto a dimensão é $n - r$.

A matriz geradora G é construída tomando os coeficientes dos polinômios da base $\{g(x), xg(x), \ldots, x^{n-r-1} g(x)\}$ como vetores linhas. Cada linha representa um deslocamento do polinômio gerador, garantindo que o código seja cíclico. □

Seja C um código cíclico com polinômio gerador $g(x)$ e polinômio verificador $h(x) = (x^n - 1)/g(x)$. Quando uma mensagem é codificada como $c(x) = f(x)g(x)$, temos a seguinte propriedade fundamental:

$$c(x)h(x) = f(x)g(x)h(x) = f(x)(x^n - 1) \equiv 0 \pmod{x^n - 1}.$$

Esta identidade estabelece que o produto de qualquer palavra-código pelo polinômio verificador resulta em zero no anel quociente $\mathbb{F}_q[x]/\langle x^n - 1\rangle$.

Expandindo o produto $c(x)h(x)$, onde:

$$c(x) = \sum_{i=0}^{n-1} c_i x^i \quad \text{e} \quad h(x) = \sum_{\ell=0}^{k} h_\ell x^\ell,$$

obtemos que o coeficiente de x^j no produto é:

$$\sum_{i=0}^{n-1} c_i h_{j-i} = 0 \quad \text{para } j = 0, 1, \ldots, n-1$$

onde os índices são calculados módulo n.

As equações anteriores podem ser expressas na forma matricial $Hc^T = 0$, onde H é a matriz de verificação de paridade:

$$H = \begin{pmatrix} h_k & h_{k-1} & \cdots & h_0 & 0 & \cdots & 0 \\ 0 & h_k & h_{k-1} & \cdots & h_0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & h_k & h_{k-1} & \cdots & h_0 \end{pmatrix}.$$

Esta matriz possui uma estrutura deslocada (banda) determinada pelos coeficientes do polinômio verificador $h(x)$.

A dimensão do código cíclico C é dada por:

$$\dim C = n - \deg(g) = \deg(h) = k$$

onde $r = \deg(g)$ é o grau do polinômio gerador e k é o grau do polinômio verificador.

Exemplo 144. *Considere o caso particular onde $n = 7$, $q = 2$ e $m = 3$. Vamos examinar as classes ciclotômicas módulo 7 sobre $\mathbb{F}_2$, determinadas pela ação da multiplicação por $q = 2$ módulo 7:*

- $C_0 = \{0\}$, *pois* $2 \cdot 0 \equiv 0 \pmod 7$.
- $C_1 = \{1, 2, 4\}$. *O conjunto é obtido multiplicando-se sucessivamente por 2:*

$$1 \cdot 2 = 2, \quad 2 \cdot 2 = 4, \quad 4 \cdot 2 = 8 \equiv 1 \pmod 7.$$

- $C_3 = \{3, 6, 5\}$. *Da mesma forma, iniciando em 3:*

$$3 \cdot 2 = 6, \quad 6 \cdot 2 = 12 \equiv 5 \pmod 7, \quad 5 \cdot 2 = 10 \equiv 3 \pmod 7.$$

Cada classe ciclotômica C_s corresponde a um polinômio minimal irredutível $M^{(s)}(x)$ sobre $\mathbb{F}_2$. O polinômio $M^{(0)}(x)$ associado a 0 é $x-1$ (que em $\mathbb{F}_2$ é $x+1$). Para as demais classes, os polinômios são calculados pelo produto dos termos $(x-\alpha^i)$ para $i \in C_s$:

$$\begin{aligned} M^{(0)}(x) &= x+1 \\ M^{(1)}(x) &= (x-\alpha)(x-\alpha^2)(x-\alpha^4) = 1+x^2+x^3 \\ M^{(3)}(x) &= (x-\alpha^3)(x-\alpha^6)(x-\alpha^5) = 1+x+x^3 \end{aligned}$$

Portanto, a fatoração completa de x^7-1 em polinômios irredutíveis sobre $\mathbb{F}_2$ é dada por:

$$x^7-1 = M^{(0)}(x)M^{(1)}(x)M^{(3)}(x) = (1+x)(1+x^2+x^3)(1+x+x^3).$$

Definição 145. *Seja $m \geq 2$ um número inteiro. O **código BCH binário que corrige dois erros** é definido como o código C de comprimento $n = 2^m-1$ cuja matriz de verificação de paridade de dimensões $2m \times (2^m-1)$ é dada por:*

$$H = \begin{pmatrix} 1 & \alpha & \alpha^2 & \cdots & \alpha^{2^m-2} \\ 1 & \alpha^3 & \alpha^6 & \cdots & \alpha^{3(2^m-2)} \end{pmatrix}$$

onde α denota um elemento primitivo do corpo de Galois $\mathbb{F}_{2^m}$.

Teorema 146. *O código BCH binário que corrige dois erros com parâmetros $n = 2^m-1$, $k = n-2m$, $d \geq 5$ e $m \geq 3$ é um código cíclico com polinômio gerador $g(x) = M^{(1)}(x)M^{(3)}(x)$, onde $\deg(M^{(1)}) = \deg(M^{(3)}) = m$. A distância mínima do código satisfaz $d \geq 5$, garantindo a capacidade de corrigir até dois erros.*

Demonstração. Para demonstrar que $g(x) = M^{(1)}(x)M^{(3)}(x)$ é efetivamente o polinômio gerador do código BCH, analisemos as condições de pertencimento ao código.

Um vetor $c = (c_0, c_1, \ldots, c_{n-1})$ pertence ao código C se e somente se satisfaz a condição de paridade:

$$Hc^T = 0$$

Esta condição matricial equivale ao seguinte sistema de equações:

$$\sum_{i=0}^{n-1} c_i\alpha^i = 0 \quad \text{e} \quad \sum_{i=0}^{n-1} c_i\alpha^{3i} = 0$$

Associando ao vetor c o polinômio $c(x) = \sum_{i=0}^{n-1} c_i x^i$, podemos reescrever estas condições na forma polinomial:

$$c(\alpha) = 0 \quad \text{e} \quad c(\alpha^3) = 0$$

Isto significa que ambos α e α^3 são raízes do polinômio $c(x)$. Consequentemente, o polinômio minimal $M^{(1)}(x)$ de α e o polinômio minimal $M^{(3)}(x)$ de α^3 devem dividir $c(x)$:

$$M^{(1)}(x) \mid c(x) \quad \text{e} \quad M^{(3)}(x) \mid c(x)$$

Como $M^{(1)}(x)$ e $M^{(3)}(x)$ são polinômios irredutíveis sobre $\mathbb{F}_2$ e distintos para $m \geq 3$, segue que seu mínimo múltiplo comum é seu produto, e portanto:

$$M^{(1)}(x)M^{(3)}(x) \mid c(x)$$

Assim, o polinômio $g(x) = M^{(1)}(x)M^{(3)}(x)$ gera exatamente o conjunto de todos os polinômios $c(x)$ que satisfazem as condições do código BCH.

Para determinar a dimensão do código, observemos que quando $m \geq 3$, ambos $M^{(1)}(x)$ e $M^{(3)}(x)$ possuem grau m. Portanto:

$$\deg(g) = \deg(M^{(1)}) + \deg(M^{(3)}) = 2m$$

A dimensão do código cíclico é então:

$$k = n - \deg(g) = (2^m - 1) - 2m$$

Esta dimensão também pode ser verificada observando que a matriz H possui $2m$ linhas linearmente independentes, ou alternativamente, considerando o grau do polinômio verificador $h(x) = (x^n - 1)/g(x)$, que é k.

A propriedade de distância mínima $d \geq 5$ será demonstrada posteriormente utilizando argumentos específicos da teoria de códigos BCH, garantindo a capacidade de correção de até dois erros. □

Determinar a distância mínima d de um código arbitrário constitui um problema computacionalmente complexo. No entanto, quando conhecemos os zeros do polinômio gerador g, é possível estabelecer limites inferiores para d. Esta abordagem fornece a base teórica para a construção dos códigos BCH com capacidade de correção de múltiplos erros.

Definição 147. *Considere um código cíclico C caracterizado por seu polinômio gerador g. Uma vez que g divide $x^n - 1$, podemos representá-lo como*

$$g(x) = \prod_{j \in K} (x - \alpha^j),$$

sendo K um conjunto formado pela combinação de certas classes ciclotômicas. Deste modo, o elemento α^i constitui uma raiz do polinômio gerador (denominado elemento nulo do

código) quando $i \in K$. Em situação contrária, α^i não é raiz de g (sendo então classificado como elemento não nulo do código).

É importante notar que os elementos não nulos do código correspondem exatamente às raízes do polinômio verificador h.

Teorema 148. *Seja C um código cíclico com polinômio gerador g. Suponha que existam inteiros $b \geq 0$ e $D \geq 1$ tais que o polinômio gerador se anula em $D-1$ potências consecutivas do elemento primitivo α, isto é:*

$$g(\alpha^b) = g(\alpha^{b+1}) = \cdots = g(\alpha^{b+D-2}) = 0.$$

Então, a distância mínima d do código satisfaz a desigualdade $d \geq D$.

Demonstração. Como o polinômio gerador $g(x)$ possui como raízes os elementos

$$\alpha^b, \alpha^{b+1}, \ldots, \alpha^{b+D-2},$$

segue que para toda palavra-código $c = (c_0, c_1, \ldots, c_{n-1}) \in C$, o polinômio associado $c(x) = \sum_{i=0}^{n-1} c_i x^i$ também se anula nestes pontos:

$$c(\alpha^b) = c(\alpha^{b+1}) = \cdots = c(\alpha^{b+D-2}) = 0.$$

Estas condições equivalem ao sistema matricial $H'c^T = 0$, onde a matriz H' é definida por:

$$H' = \begin{pmatrix} 1 & \alpha^b & \alpha^{2b} & \cdots & \alpha^{(n-1)b} \\ 1 & \alpha^{b+1} & \alpha^{2(b+1)} & \cdots & \alpha^{(n-1)(b+1)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \alpha^{b+D-2} & \alpha^{2(b+D-2)} & \cdots & \alpha^{(n-1)(b+D-2)} \end{pmatrix}$$

Para demonstrar que $d \geq D$, vamos provar que quaisquer $D-1$ colunas de H' são linearmente independentes. Suponhamos, por contradição, que exista uma palavra-código não nula c com peso $\ell \leq D-1$. Sejam $a_1, a_2, \ldots, a_\ell$ as posições onde $c_i \neq 0$.

Do sistema $H'c^T = 0$, extraímos as colunas correspondentes às posições não nulas, obtendo o subsistema:

$$\begin{pmatrix} \alpha^{a_1 b} & \alpha^{a_2 b} & \cdots & \alpha^{a_\ell b} \\ \alpha^{a_1(b+1)} & \alpha^{a_2(b+1)} & \cdots & \alpha^{a_\ell(b+1)} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha^{a_1(b+D-2)} & \alpha^{a_2(b+D-2)} & \cdots & \alpha^{a_\ell(b+D-2)} \end{pmatrix} \begin{pmatrix} c_{a_1} \\ c_{a_2} \\ \vdots \\ c_{a_\ell} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

Este é um sistema homogêneo com $D-1$ equações e $\ell \leq D-1$ incógnitas. Para que exista solução não trivial, a matriz do sistema deve ser singular. No entanto, vamos demonstrar que tal matriz possui determinante não nulo.

Fatorando $\alpha^{a_j b}$ de cada coluna j, obtemos:

$$\det = \left(\prod_{j=1}^{\ell} \alpha^{a_j b}\right) \begin{vmatrix} 1 & 1 & \cdots & 1 \\ \alpha^{a_1} & \alpha^{a_2} & \cdots & \alpha^{a_\ell} \\ \alpha^{2a_1} & \alpha^{2a_2} & \cdots & \alpha^{2a_\ell} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha^{(D-2)a_1} & \alpha^{(D-2)a_2} & \cdots & \alpha^{(D-2)a_\ell} \end{vmatrix}$$

O determinante resultante é de Vandermonde e pode ser expresso como:

$$\prod_{1 \le i < j \le \ell} (\alpha^{a_j} - \alpha^{a_i})$$

Como α é um elemento primitivo e os índices a_i são distintos, temos $\alpha^{a_i} \neq \alpha^{a_j}$ para $i \neq j$, garantindo que o determinante seja diferente de zero. Portanto, a única solução do sistema homogêneo é a trivial, contradizendo a existência de uma palavra-código não nula com peso menor que D.

Concluímos assim que $d \geq D$, completando a demonstração. □

Definição 149. *Um código cíclico de comprimento n definido sobre o corpo finito $\mathbb{F}_q$ é denominado* ***código BCH com distância de projeto*** *D quando existem inteiros $b \geq 0$ e $D \geq 1$ para os quais o polinômio gerador é determinado por:*

$$g(x) = mmc\left(M^{(b)}(x), M^{(b+1)}(x), \ldots, M^{(b+D-2)}(x)\right)$$

Equivalentemente, $g(x)$ representa o polinômio mônico de menor grau sobre $\mathbb{F}_q$ que possui como raízes os elementos $\alpha^b, \alpha^{b+1}, \ldots, \alpha^{b+D-2}$. Consequentemente, um vetor $c = (c_0, c_1, \ldots, c_{n-1})$ pertence ao código C se e somente se satisfaz as condições:

$$c(\alpha^b) = c(\alpha^{b+1}) = \cdots = c(\alpha^{b+D-2}) = 0$$

Teorema 150. *Um código BCH sobre $\mathbb{F}_q$ de comprimento n e distância de projeto D possui distância mínima $d \geq D$ e dimensão $k \geq n - m(D-1)$.*

Demonstração. A demonstração divide-se em duas partes principais:

Parte 1: Demonstração de que $d \geq D$

Seja C um código BCH com polinômio gerador $g(x)$ definido sobre $\mathbb{F}_q$. Por definição, $g(x)$ é o polinômio minimal de menor grau que possui como raízes os elementos $\alpha^b, \alpha^{b+1}, \ldots, \alpha^{b+D-2}$, onde α é uma raiz n-ésima primitiva da unidade.

Para qualquer palavra-código $c = (c_0, c_1, \ldots, c_{n-1}) \in C$, temos:

$$c(\alpha^b) = c(\alpha^{b+1}) = \cdots = c(\alpha^{b+D-2}) = 0.$$

Estas condições equivalem ao sistema matricial $Hc^T = 0$, onde a matriz de verificação de paridade é:

$$H = \begin{pmatrix} 1 & \alpha^b & \alpha^{2b} & \cdots & \alpha^{(n-1)b} \\ 1 & \alpha^{b+1} & \alpha^{2(b+1)} & \cdots & \alpha^{(n-1)(b+1)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \alpha^{b+D-2} & \alpha^{2(b+D-2)} & \cdots & \alpha^{(n-1)(b+D-2)} \end{pmatrix}.$$

Cada elemento α^{ij} nesta matriz pertence a $\mathbb{F}_{q^m}$, e pode ser representado como um vetor coluna de m componentes em $\mathbb{F}_q$. Assim, a matriz H expandida tem dimensões $(D-1)m \times n$.

Pelo Teorema 148 ,que quaisquer $D-1$ colunas de H são linearmente independentes. Consequentemente, o código possui distância mínima $d \geq D$.

Parte 2: Demonstração de que $k \geq n - m(D-1)$

O polinômio gerador $g(x)$ é dado por:

$$g(x) = \text{mmc}\{M^{(b)}(x), M^{(b+1)}(x), \ldots, M^{(b+D-2)}(x)\}$$

onde $M^{(i)}(x)$ é o polinômio minimal de α^i sobre $\mathbb{F}_q$.

Cada polinômio minimal $M^{(i)}(x)$ tem grau no máximo m, pois $\alpha^i \in \mathbb{F}_{q^m}$ e o grau do polinômio minimal é igual ao tamanho da classe ciclotômica contendo i, que é no máximo m.

Como $g(x)$ é o mínimo múltiplo comum de $D-1$ polinômios, cada um com grau $\leq m$, temos:

$$\deg(g) \leq m(D-1)$$

A dimensão do código cíclico C é dada por:

$$k = n - \deg(g)$$

Portanto:

$$k = n - \deg(g) \geq n - m(D-1)$$

Isto completa a demonstração das desigualdades $d \geq D$ e $k \geq n - m(D-1)$. □

Exemplo 151. *Considere o caso em que $q = 2$ e $n = 23$ (que não é um comprimento primitivo). Iniciamos determinando as classes laterais ciclotômicas módulo 23:*

$$C_0 = \{0\},$$
$$C_1 = \{1, 2, 4, 8, 16, 9, 18, 13, 3, 6, 12\},$$
$$C_5 = \{5, 10, 20, 17, 11, 22, 21, 19, 15, 7, 14\}.$$

Dessas classes, obtemos as seguintes informações:

$$|C_1| = 11 \quad \Rightarrow \quad \deg(M^{(1)}) = 11.$$
$$|C_5| = 11 \quad \Rightarrow \quad \deg(M^{(5)}) = 11.$$

A decomposição do polinômio $x^{23} - 1$ *em fatores irredutíveis é dada por:*

$$x^{23} - 1 = (x - 1) \cdot M^{(1)}(x) \cdot M^{(5)}(x),$$

onde os polinômios minimais são:

$$M^{(1)}(x) = x^{11} + x^9 + x^7 + x^6 + x^5 + x + 1.$$
$$M^{(5)}(x) = x^{11} + x^{10} + x^6 + x^5 + x^4 + x^2 + 1.$$

A tabela abaixo resume os parâmetros dos códigos BCH obtidos:

Distância de projeto	*Polinômio gerador*	*Dimensão* $n - \deg(g)$	*Distância real*
1	*1*	*23*	*1*
3 ou 5	$M^{(1)}$	*12*	*7*
7, 9, ..., 23	$M^{(1)}M^{(5)}$	*1*	*23*

Referências

BREšAR, M. *Undergraduate Algebra: A Unified Approach*. [S.l.]: Springer, 2019. Citado 3 vezes nas páginas 9, 13 e 16.

LIDL, R.; NIEDERREITER, H. *Finite Fields*. [S.l.]: Cambridge University Press, 1997. Citado 3 vezes nas páginas 9, 10 e 43.

LIDL, R.; PILZ, G. *Applied Abstract Algebra*. 2. ed. [S.l.]: Springer, 1997. Citado 3 vezes nas páginas 9, 45 e 65.

PANARIO, A. M. M. e D. *Tópicos de Corpos Finitos com Aplicações em Criptografia e Teoria de Códigos*. [S.l.]: IMPA, 2007. Citado 3 vezes nas páginas 9, 45 e 74.